

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 09:56:46 ; Search time 45 Seconds  
(without alignments)  
2902.767 Million cell updates/sec

Title: US-09-978-318B-1  
Perfect score: 2203  
Sequence: 1 MAALRYAGLDTDSEDELLP.....RTLWALSERLIQERLGSQSG 414

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTRMBL\_25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2203	100.0	414	4	Q9NZC7
2	2191	99.5	414	4	Q9NRF5
3	2071	94.0	414	11	Q91WL8
4	2067	93.8	414	11	Q9JLF5
5	1869	84.8	363	4	Q9NRF6
6	1785	81.0	354	11	Q9D2B3
7	1777	80.7	367	11	Q9D339
8	1637.5	74.3	412	13	Q803A8
9	1157	52.5	234	4	Q9BTT8
10	1149	52.2	242	11	Q920T2
11	1023	46.4	409	5	Q9VLU5
12	942	42.8	191	4	Q9NRK1
13	927	42.1	189	4	Q9NPC9
14	750.5	34.1	213	4	Q9KMK3
15	746.5	33.9	311	4	Q9GRF2
16	740	33.6	158	11	Q8C8J6

17	541.5	24.6	330	5	Q960C7
18	539	24.5	370	5	Q9UAW2
19	538.5	24.4	329	5	Q19062
20	538.5	24.4	331	5	Q9V4P9
21	537.5	24.4	330	5	Q9V4Q3
22	537	24.4	321	10	Q9LGI8
23	528	24.0	319	5	Q95QN9
24	520.5	23.6	320	10	Q8L9T6
25	520.5	23.6	325	5	Q8MZG9
26	514.5	23.4	316	10	Q8HIQ6
27	510.5	23.2	300	5	Q9V4Q2
28	509.5	23.1	296	5	Q8MKN1
29	506.5	23.0	293	11	Q9RIR9
30	506.5	23.0	300	11	Q9RIR8
31	497	22.6	317	10	Q9LDY7
32	497	22.6	321	10	Q80924
33	495	22.5	331	10	Q8RWJ2
34	493	22.4	319	10	Q94LU0
35	493	22.4	319	10	Q7XCH3
36	490.5	22.3	370	4	Q12953
37	489	22.2	331	10	Q8LCE7
38	485.5	22.0	309	13	Q7T348
39	485.5	22.0	406	5	Q9W404
40	478	21.7	379	5	Q45680
41	476.5	21.6	350	10	Q9LZ48
42	472	21.4	336	16	Q9RR99
43	469	21.3	314	10	Q7XMT5
44	469	21.3	320	10	P81259
45	447.5	20.3	294	10	O81739

## ALIGNMENTS

## RESULT 1

Q9NZC7	PRELIMINARY;	PRT;	414 AA.
ID	Q9NZC7	PRELIMINARY;	PRT;
AC	Q9NZC7	PRELIMINARY;	PRT;
DT	01-OCT-2000 (TrEMBLrel. 15, Created)		
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	WW domain-containing protein WWOX (WW domain-containing oxidoreductase isoform FORII).		
DE	isoform FORII.		
GN	WWOX.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]_TaxID=9606;		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=20246348; PubMed=10786576;		
RA	Bednarek A.K., Laflin K.J., Daniel R.L., Liao Q., Hawkins K.A.,		
RA	Aldaz C.M.;		
RT	"WWOX, a novel WW domain-containing protein mapping to human		
RT	chromosome 16q23.3-24.1, a region frequently affected in breast		
RT	cancer.";		
RL	Cancer Res. 60:2140-2145 (2000).		
RN	[2]		
RN	SEQUENCE FROM N.A.		
RX	PubMed=11572989;		
RA	Paige A.J.W., Taylor K.J., Taylor C., Hallier S.G., Farrington S.,		
RA	Scott D., Porteous D.J., Smyth J.F., Gabra H., Watson J.E.V.;		
RT	"WWOX: A candidate tumor suppressor gene involved in multiple tumor		
RT	types.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 98:11417-11422 (2001).		
CC	- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES		
CC	(SDR) FAMILY.		
DR	EMBL; AF211943; AAF27049.1; -		
DR	EMBL; AF325432; AAL05449.1; -		
DR	EMBL; AF325433; AAL05449.1; JOINED.		
DR	EMBL; AF325424; AAL05449.1; JOINED.		
DR	EMBL; AF325425; AAL05449.1; JOINED.		
DR	EMBL; AF325426; AAL05449.1; JOINED.		

DR EMBL; AF325427; AAL05449.1; JOINED.  
 DR EMBL; AF325428; AAL05449.1; JOINED.  
 DR EMBL; AF325430; AAL05449.1; JOINED.  
 DR EMBL; AF325431; AAL05449.1; JOINED.  
 DR HSSP; Q13526; IPIN.  
 DR GO; GO:0005489; F:electron transporter activity; TAS.  
 DR GO; GO:0005515; F:protein binding; TAS.  
 DR GO; GO:0008202; P:steroid metabolism; TAS.  
 DR InterPro; IPR002198; ADH short.  
 DR InterPro; IPR002198; ADH short.  
 DR Pfam; PF00106; adh\_short; 1.  
 DR Pfam; PF00397; WW; 2.  
 DR SMART; SM00456; WW; 2.  
 DR PROSITE; PS01159; WW DOMAIN 1; 2.  
 DR PROSITE; PS00200; WW DOMAIN 2; 2.  
 DR Oxidoreductase.  
 KW Oxidoreductase.  
 SQ SEQUENCE 414 AA; 46676 MW; E4D9A649B6CB05DF CRC64;

Query Match 100.0%; Score 2203; DB 4; Length 414;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-176;  
 Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAALRYAGLDDTSDSELPGEERTTKDGVVYANHTTEKTQWEHPKTKRKRKRVAGDLP 60  
 DB 1 MAALRYAGLDDTSDSELPGEERTTKDGVVYANHTTEKTQWEHPKTKRKRKRVAGDLP 60  
 QY 61 YGWEQETDENGQVFFVDHINKRTTYLDPRLAFTVDDNPTKPTRQRYDGSSTTAMEILQGR 120  
 DB 61 YGWEQETDENGQVFFVDHINKRTTYLDPRLAFTVDDNPTKPTRQRYDGSSTTAMEILQGR 120  
 QY 121 DFTGKVVVVTGANGSIGGFETAKSFALHGAHVILACRNMARASEAVSRILEEWHKAKVEAM 180  
 DB 121 DFTGKVVVVTGANGSIGGFETAKSFALHGAHVILACRNMARASEAVSRILEEWHKAKVEAM 180  
 QY 181 TLDLALLRSVOHFAEAFKAKNVLPLVLCNAATFALPWSLTKDGLTTFQVNLGHFYL 240  
 DB 181 TLDLALLRSVOHFAEAFKAKNVLPLVLCNAATFALPWSLTKDGLTTFQVNLGHFYL 240  
 QY 241 QLLQDVLCSPARVIVVSSSHRFTDINDSLGKLDPSRLSPKNDYWLAMLAYNRSKLCN 300  
 DB 241 QLLQDVLCSPARVIVVSSSHRFTDINDSLGKLDPSRLSPKNDYWLAMLAYNRSKLCN 300  
 QY 301 ILFNSNELHRLSPRGVTSNAVHPGNMYSNIHRSWWVYLLFTLARPFTKSMQCGAATT 360  
 DB 301 ILFNSNELHRLSPRGVTSNAVHPGNMYSNIHRSWWVYLLFTLARPFTKSMQCGAATT 360  
 QY 361 YCAAVPELEGGMVFNCCRCMPSPQAQSEETARTLWALSRLIQLERLGSQSG 414  
 DB 361 YCAAVPELEGGMVFNCCRCMPSPQAQSEETARTLWALSRLIQLERLGSQSG 414

RESULT 2  
 Q9NRF5 PRELIMINARY; PRT; 414 AA.  
 ID Q9NRF5  
 AC Q9NRF5  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DE 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
 DE For II protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20320695; PubMed=10861292;  
 RA Ried K., Finniss M., Hobson L., Mangelsdorf M., Dayan S.,  
 RA Nancarrow J.K., Woolatt E., Kremmidiotis G., Gardner A., Venter D.,  
 RA Baker E., Richards R.I.,  
 RT "Common chromosomal fragile site FRA16D sequence: Identification of the  
 RT POR gene spanning FRA16D and homozygous deletions and translocation  
 RT breakpoints in cancer cells.";  
 RT Cancer 9:122-127 (1994)

CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
 CC (SDR) FAMILY.  
 CC EMBL; AF227527; AAF82054.1; --  
 DR HSSP; Q13526; IPIN.  
 DR Genew; HGNC:12799; WWOX.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR002198; ADH short.  
 DR InterPro; IPR002198; ADH short.  
 DR Pfam; PF00106; adh\_short; 1.  
 DR Pfam; PF00397; WW; 2.  
 DR SMART; SM00456; WW; 2.  
 DR PROSITE; PS01159; WW DOMAIN 1; 2.  
 DR PROSITE; PS00200; WW DOMAIN 2; 2.  
 DR Oxidoreductase.  
 KW Oxidoreductase.  
 SQ SEQUENCE 414 AA; 46680 MW; 04D9B358F87F0E75 CRC64;

Query Match 99.5%; Score 2191; DB 4; Length 414;  
 Best Local Similarity 99.5%; Pred. No. 1.6e-175;  
 Matches 412; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAALRYAGLDDTSDSELPGEERTTKDGVVYANHTTEKTQWEHPKTKRKRKRVAGDLP 60  
 DB 1 MAALRYAGLDDTSDSELPGEERTTKDGVVYANHTTEKTQWEHPKTKRKRKRVAGDLP 60  
 QY 61 YGWEQETDENGQVFFVDHINKRTTYLDPRLAFTVDDNPTKPTRQRYDGSSTTAMEILQGR 120  
 DB 61 YGWEQETDENGQVFFVDHINKRTTYLDPRLAFTVDDNPTKPTRQRYDGSSTTAMEILQGR 120  
 QY 121 DFTGKVVVVTGANGSIGGFETAKSFALHGAHVILACRNMARASEAVSRILEEWHKAKVEAM 180  
 DB 121 DFTGKVVVVTGANGSIGGFETAKSFALHGAHVILACRNMARASEAVSRILEEWHKAKVEAM 180  
 QY 181 TLDLALLRSVOHFAEAFKAKNVLPLVLCNAATFALPWSLTKDGLTTFQVNLGHFYL 240  
 DB 181 TLDLALLRSVOHFAEAFKAKNVLPLVLCNAATFALPWSLTKDGLTTFQVNLGHFYL 240  
 QY 241 QLLQDVLCSPARVIVVSSSHRFTDINDSLGKLDPSRLSPKNDYWLAMLAYNRSKLCN 300  
 DB 241 QLLQDVLCSPARVIVVSSSHRFTDINDSLGKLDPSRLSPKNDYWLAMLAYNRSKLCN 300  
 QY 301 ILFNSNELHRLSPRGVTSNAVHPGNMYSNIHRSWWVYLLFTLARPFTKSMQCGAATT 360  
 DB 301 ILFNSNELHRLSPRGVTSNAVHPGNMYSNIHRSWWVYLLFTLARPFTKSMQCGAATT 360  
 QY 361 YCAAVPELEGGMVFNCCRCMPSPQAQSEETARTLWALSRLIQLERLGSQSG 414  
 DB 361 YCAAVPELEGGMVFNCCRCMPSPQAQSEETARTLWALSRLIQLERLGSQSG 414

RESULT 3  
 Q91WL8 PRELIMINARY; PRT; 414 AA.  
 ID Q91WL8  
 AC Q91WL8  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DE 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
 DE WW domain-containing oxidoreductase.  
 GN WWOX.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Kidney;  
 RA Strausberg R.;  
 RN Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6J; TISSUE=Mullerian duct includes surrounding region;  
 RC MEDLINE=22354683; PubMed=12466851;  
 RX

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2004, 21:49:29 ; Search time 8933 Seconds  
(without alignments)  
11109.321 Million cell updates/sec

Title: US-09-978-318B-2  
Perfect score: 2264  
Sequence: 1 gcagtcgcagcgagcg.....aaaaaaaaaaaaaaaaaaaaa 2264

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rtd.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2264	100.0	2264	9	AF211943	AF211943 Homo sapi
2	2253.4	99.5	2256	6	AX877089	AX877089 Sequence
3	2253.4	99.5	2256	6	BD012189	BD012189 A novel g
4	2253.4	99.5	2256	6	BD156478	BD156478 Primer fo
5	2253.4	99.5	2256	6	AK027626	AK027626 Homo sapi
6	2213.4	97.8	2219	9	AF227527	AF227527 Homo sapi
7	1372.4	60.6	1475	9	HS013395	UI3395 Human oxido
8	1174	51.9	1732	9	AF395123	AF395123 Homo sapi
9	1137.4	50.2	1225	9	AF227526	AF227526 Homo sapi
10	1096.6	48.4	2269	10	BC014716	BC014716 Mus muscu
11	1092.8	48.3	1669	9	BC003184	BC003184 Homo sapi
12	1084	47.9	1625	9	AF395124	AF395124 Homo sapi
13	1077	47.6	2197	10	AF187014	AF187014 Mus muscu
14	1062.2	46.9	168083	9	AC092376	AC092376 Homo sapi
15	1052.6	46.5	161988	2	AC009129	AC009129 Homo sapi
16	691.6	30.5	768	6	AX868436	AX868436 Sequence
17	691.6	30.5	768	6	BD148498	BD148498 Primer fo
18	624.8	27.6	1914	5	BC044560	BC044560 Danilo rer
19	617.6	27.3	74586	2	AC060793	AC060793 Homo sapi
20	616.4	27.2	711	9	AF227528	AF227528 Homo sapi
21	567.2	25.1	729	9	AF187015	AF187015 Homo sapi
22	518.2	22.9	705	12	BT007739	BT007739 Synthetic
23	516.6	22.8	705	9	BT007445	BT007445 Homo sapi
24	483.6	21.4	541	6	AX873488	AX873488 Sequence
25	483.6	21.4	541	6	BD153550	BD153550 Primer fo
26	433.6	19.2	505	9	F325423S11	F325423 Homo sapi
27	416.8	18.4	75852	2	AC021253	AC021253 Homo sapi
28	342.4	15.1	1763	3	AK115111	AK115111 Ciona int
29	280.6	12.4	298	11	G15863	G15863 human STS C
30	270.4	11.9	1460	3	AY119574	AY119574 Drosophil
31	266	11.7	457	9	F325423S09	F325423 Homo sapi
32	266	11.7	147101	9	AC109134	AC109134 Homo sapi
33	266	11.7	270752	9	AF217491S3	AF217490 Homo sapi
34	249.4	11.0	258912	2	AC129777	AC129777 Mus muscu
35	237.8	10.5	254	11	G14545	G14545 human STS S
36	235.2	10.4	74586	2	AC060793	AC060793 Homo sapi
37	232	10.2	431	9	F325423S01	F325423 Homo sapi
38	232	10.2	1440	9	AF212843	AF212843 Homo sapi
39	232	10.2	167015	9	AC079414	AC079414 Homo sapi
40	232	10.2	167195	9	AC009044	AC009044 Homo sapi
41	211.6	9.3	220157	2	AC098432	AC098432 Rattus no
42	211.6	9.3	22465	2	AC130071	AC130071 Rattus no
43	211.6	9.3	255094	2	AC098195	AC098195 Rattus no
44	211	9.3	524	9	AF227529	AF227529 Homo sapi
45	206.8	9.1	778	10	AY046553S3	AY046555 Mus muscu

ALIGNMENTS

RESULT 1  
AF211943  
LOCUS AF211943 2264 bp mRNA linear PRI 05-MAY-2000  
DEFINITION Homo sapiens WW domain-containing protein WWOX mRNA, complete cds.  
ACCESSION AF211943  
VERSION AF211943.1 GI:6729682  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2264)  
Bednarek A.K., Laflin, K.J., Daniel, R.L., Liao, Q., Hawkins, K.A. and  
Aldaz, C.M.  
AUTHORS WWOX, a novel WW domain-containing protein mapping to human  
TITLE

chromosome 16q23.3-24.1, a region frequently affected in breast

Cancer Res. 60 (8), 2140-2145 (2000)  
 20246348  
 10786676  
 2 (bases 1 to 2264)  
 Bednarek, A.K. and Aldaz, C.M.  
 Direct Submission  
 Submitted (06-DEC-1999) Carcinogenesis, University of Texas M.D.  
 Anderson Cancer Center, Science Park-Research Division, Park Road  
 1C POB389, Smithville, TX 78957, USA

Location/Qualifiers  
 1. .2264  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="16"  
 /map="16q23.3-q24.1"  
 1. .125  
 126. .1370  
 /note="putative hydroxysteroid dehydrogenase"  
 /codon\_start=1  
 /product="WW domain-containing protein WWOX"  
 /protein\_id="AA27049.1"  
 /db\_xref="GI:6729683"  
 /translation="MAALRYAGLDDTDSDELPPWHEERTTKDGVVYANHTREKQW  
 EHPKTRKRVAGDPLPYGMEQETDENGQVFVPHINKRTYLDLPRIATVDDNPTPT  
 TRQYDGSITAMETLOQRDFTGVVVTGANSIGIFETAKSFALHGHVILACRNWAR  
 ASERVLLEHMKVQAMTLDLALRSVQHPAEAFKAKNVLPLVLCNAATFALPFW  
 SLTDGLGTFQVNLGHFVLQDLVLCRSAPARVIVVSSSHRFTDINDSLGLD  
 FSLSPKTDYWMALNRSKCLNLFSELNRLRLSPRGVTSNAPVHGNMWNYSIHR  
 WYVTLFTLARPFTSMQQAATVYCAAVPELGLGLGMFNNCCRCMPSPQAQSEE  
 TARTLWLSERLIQERLGSQSG"  
 1371. .2264  
 2217. .2222

3'UTR polyA\_signal ORIGIN

Query Match 100.0%; Score 2264; DB 9; Length 2264;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 GCAGTGGCAGCGCTGAGCGGTGGGGCCCGCGCGCGGTCTCGTTTGGAGCGGGAG 60  
 1 GCAGTGGCAGCGCTGAGCGGTGGGGCCCGCGCGCGGTCTCGTTTGGAGCGGGAG 60  
 61 TGAGTTCCTGACGAGTGGACCGCGCAGCGCGCGATAGGGGGCCAGGTGCTCCACAGT 120  
 61 TGAGTTCCTGACGAGTGGACCGCGCAGCGCGCGATAGGGGGCCAGGTGCTCCACAGT 120  
 121 CAGCCATGGCAGCGCTGCGCTACGCGCGCGCTGGACGACGACGACGACGACGACGACG 180  
 121 CAGCCATGGCAGCGCTGCGCTACGCGCGCGCTGGACGACGACGACGACGACGACGACG 180  
 181 CTCGGGCTGGGAGGAGAGACCAACGAGCGCTGGGTCTTACTACGCGCAATCAACCG 240  
 181 CTCGGGCTGGGAGGAGAGACCAACGAGCGCTGGGTCTTACTACGCGCAATCAACCG 240  
 241 AGGAGAGAGACTCAGTGGGAGACATCCAAAATCTGGAAAAAGAAAACGAGTGGCAGGAGT 300  
 241 AGGAGAGAGACTCAGTGGGAGACATCCAAAATCTGGAAAAAGAAAACGAGTGGCAGGAGT 300  
 301 TGCCATACGATGGGAAACAAGAACTGATGAGAACGACGACGACGACGACGACGACGACG 360  
 301 TGCCATACGATGGGAAACAAGAACTGATGAGAACGACGACGACGACGACGACGACGACG 360  
 361 TAAATAAAGAACCACTTCTTGACCAACGACGACGACGACGACGACGACGACGACGACG 420  
 361 TAAATAAAGAACCACTTCTTGACCAACGACGACGACGACGACGACGACGACGACGACG 420  
 421 CCAAGCCAAACCAACCGGCAAGATACGACGACGACGACGACGACGACGACGACGACG 480  
 421 CCAAGCCAAACCAACCGGCAAGATACGACGACGACGACGACGACGACGACGACGACG 480

481 GCCGGGATTTCACTGGCAAGTGGTGTGGTCACTGGAGCTAAATTCAGGAATAGGGTTGG 540  
 481 GCCGGGATTTCACTGGCAAGTGGTGTGGTCACTGGAGCTAAATTCAGGAATAGGGTTGG 540  
 541 AAACCGCCAGTCTTTTGGCCCTCCATGGTGCAATGTGATCTTGGCCTGCGAGGAACATGG 600  
 541 AAACCGCCAGTCTTTTGGCCCTCCATGGTGCAATGTGATCTTGGCCTGCGAGGAACATGG 600  
 601 CAAAGCGGAGTGAAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 660  
 601 CAAAGCGGAGTGAAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 660  
 661 CAATGACCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
 661 CAATGACCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
 721 CCAAGAAATGCTCTTTCATGTGCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
 721 CCAAGAAATGCTCTTTCATGTGCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
 781 GTCTCAACAAAGATGGCTGGAGACCACTTTCATGTGCTGCTGCTGCTGCTGCTGCTG 840  
 781 GTCTCAACAAAGATGGCTGGAGACCACTTTCATGTGCTGCTGCTGCTGCTGCTGCTG 840  
 841 TTGTCCAGCTCTTCCAGGATGTTTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
 841 TTGTCCAGCTCTTCCAGGATGTTTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
 901 CCTCAGAGTCCCATCGATTTACAGATATTAACGACTCTTGGGAAAACTGGACTTCAGTC 960  
 901 CCTCAGAGTCCCATCGATTTACAGATATTAACGACTCTTGGGAAAACTGGACTTCAGTC 960  
 961 GCCTCTCTCCAAACAAAAACGACTATTTGGGCGATGCTGGCTTATTAACAGTCCCAAGCT 1020  
 961 GCCTCTCTCCAAACAAAAACGACTATTTGGGCGATGCTGGCTTATTAACAGTCCCAAGCT 1020  
 1021 GCAACATCTCTTCTCCAAACGAGTGCACCGTCCCTCTCCCGACGCGGGGTCACTGCGA 1080  
 1021 GCAACATCTCTTCTCCAAACGAGTGCACCGTCCCTCTCCCGACGCGGGGTCACTGCGA 1080  
 1081 ACAGAGTGCATCTCGAAAAATATGATGATGATGATGATGATGATGATGATGATGATG 1140  
 1081 ACAGAGTGCATCTCGAAAAATATGATGATGATGATGATGATGATGATGATGATGATG 1140  
 1141 CACTGCTGTTTACCTTGGCGAGGCTTTCACCAAGTCCATGCAACAGGAGCTGCACCA 1200  
 1141 CACTGCTGTTTACCTTGGCGAGGCTTTCACCAAGTCCATGCAACAGGAGCTGCACCA 1200  
 1201 CGGTGTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260  
 1201 CGGTGTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260  
 1261 GCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
 1261 GCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
 1321 CGCTCAGCGAGAGGTGATCCAAAGACGCTTGGCAGCGAGTCCGCTTAAGTGGAGCTCA 1380  
 1321 CGCTCAGCGAGAGGTGATCCAAAGACGCTTGGCAGCGAGTCCGCTTAAGTGGAGCTCA 1380  
 1381 GAGCGATGGGACACACACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440  
 1381 GAGCGATGGGACACACACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440  
 1441 GCCCTTCCAAATGCTCCCTCCAAACAGATCCGCAAGATGAAAGAAATTAAGAGCAGTCA 1500  
 1441 GCCCTTCCAAATGCTCCCTCCAAACAGATCCGCAAGATGAAAGAAATTAAGAGCAGTCA 1500  
 1501 CAACAGAGTGAATAATCTTAAAGTACCAATGGGAGCAGGGAATCTCTGGGTAAGATATC 1560  
 1501 CAACAGAGTGAATAATCTTAAAGTACCAATGGGAGCAGGGAATCTCTGGGTAAGATATC 1560  
 1561 ACTTTTCTGGGGCTGGGCTAGGCTAGTCTCTTTTCTGTTCTGTTGCTGCTGTTTGA 1620

```
Db 1561 ACTTTCTGGGGCTGGGCTAGCATAGTCTCTTTGCTTTCTGGTGGCTGTTGAA 1620
Qy 1621 AGTAAACCTGCTGGTGTGTAGTTCCTGATTCCTCGAGAGACACAGCAATCTC 1680
Db 1621 AGTAAACCTGCTGGTGTGTAGTTCCTGATTCCTCGAGAGACACAGCAATCTC 1680
Qy 1681 TTTCTTTTACTGTTATAGATAGACCTGAGTCCCTGCTCCATCCAGTACCAACAGG 1740
Db 1681 TTTCTTTTACTGTTATAGATAGACCTGAGTCCCTGCTCCATCCAGTACCAACAGG 1740
Qy 1741 CCACCACTGCAGCCGGGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1800
Db 1741 CCACCACTGCAGCCGGGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1800
Qy 1801 TGCTCTCTGCTGATTCAGAGGATATTTGTTTCAATTCCTGACCAAGACTGAGC 1860
Db 1801 TGCTCTCTGCTGATTCAGAGGATATTTGTTTCAATTCCTGACCAAGACTGAGC 1860
Qy 1861 CAGCTTAGCAATCTGCTGGGAGACAAATCTCAGAACCTTTGTCAGCCAGTGAGTAC 1920
Db 1861 CAGCTTAGCAATCTGCTGGGAGACAAATCTCAGAACCTTTGTCAGCCAGTGAGTAC 1920
Qy 1921 AGTGACACCGAGGAGTAGATACGAGAACTACAGGTGTCAGAGTCTGTCATAG 1980
Db 1921 AGTGACACCGAGGAGTAGATACGAGAACTACAGGTGTCAGAGTCTGTCATAG 1980
Qy 1981 ACTCTCTTCTGATGCTGCAAAAAATTTCTTTAGAGATTATACAAATTTTCAAATCA 2040
Db 1981 ACTCTCTTCTGATGCTGCAAAAAATTTCTTTAGAGATTATACAAATTTTCAAATCA 2040
Qy 2041 TTCCTTAGATACCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2100
Db 2041 TTCCTTAGATACCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2100
Qy 2101 TTTTGGGGGCGAGAGATTAAGCTTAGTTAATCCCTTTCTGCTGCTGCTGCTGCTGCT 2160
Db 2101 TTTTGGGGGCGAGAGATTAAGCTTAGTTAATCCCTTTCTGCTGCTGCTGCTGCTGCT 2160
Qy 2161 TTTCTCTGCTTTCACATTTGATTTAAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
Db 2161 TTTCTCTGCTTTCACATTTGATTTAAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
Qy 2221 AAAGAACTGCTTGAATATCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 2264
Db 2221 AAAGAACTGCTTGAATATCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 2264

RESULT 2
AX877089 2256 bp DNA linear PAT 17-DEC-2003
LOCUS
DEFINITION Sequence 11994 from Patent EP1074617.
ACCESSION AX877089
VERSION AX877089.1 GI:40031825
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Negai, K. and Otsuki, T.
Primers for synthesizing full-length cDNA and their use
Patent: EP 1074617-A 11994 07-FEB-2001;
Research Association for Biotechnology (JP)
FEATURES
Location/Qualifiers
source
1..2256
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
125..1369
/note="unnamed protein product"
/codon_start=1
CDS
```

```
/protein_id="CAE89566.1"
/db_xref="GI:40031826"
/translation="MAALRVAGLDDTDESELPFGWEERTTKDQWVYVYANHEBKTOW
EHPDTGKRVAGDLPGWEOTDENGQVFDHINKRTVLDPLRAHVLACRNMAR
TRQYDGSITAMEILOPQDFDKVVTGANSIGFETAKSFAHGHARVILACRNMAR
ASEAVSRILEEMHAKAVEAMTDLALLRSVOHFAPBAKAKVPLHVLVCRNATFALPW
SLTDGLETTTFQVNHGHFYLQDLQVLCBSAPARVTVVSSSEHRTFDINDSKLD
FSRLSPKNDYWMALVNRSLKILFNSLHRELSRPGVTSNAPHPNMMYSNIHRS
WVVVTLFLPTLARPTKSMQOGAAATTVCAVPELEGLGMYFNCCRCMPFPAQSEE
TARTLWALSERLIQERLGSQS"
ORIGIN
Query Match 99.5%; Score 2253.4; DB 6; Length 2256;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 CAGTCCGAGGCGTGAAGCGTCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 61
Db 1 CAGTCCGAGGCGTGAAGCGTCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60
Qy 62 GAGTTTCTCAGCGAGTGGACCCGCGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 121
Db 61 GAGTTTCTCAGCGAGTGGACCCGCGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
Qy 122 AGCCATGGCAGCGCTGCGCTACGCGGGGCTGACGACGACGACGACGACGACGACGACG 181
Db 121 AGCCATGGCAGCGCTGCGCTACGCGGGGCTGACGACGACGACGACGACGACGACGACG 180
Qy 182 TCCGGGCTGGGAGGAGAGAACCAACAGGACCGGCTGGGTTTACTAGCCCAATCACACCGA 241
Db 181 TCCGGGCTGGGAGGAGAGAACCAACAGGACCGGCTGGGTTTACTAGCCCAATCACACCGA 240
Qy 242 GAGAGAGACTCAGTGGGAACATCCAAAACTGMAAAAGAAAAACGAGTGGCAGGAGATT 301
Db 241 GAGAGAGACTCAGTGGGAACATCCAAAACTGMAAAAGAAAAACGAGTGGCAGGAGATT 300
Qy 302 GCCATACCGATGGGAACAAAGAACTGATGAGAACCGCAAGTGTGTTTGTTCACCATAT 361
Db 301 GCCATACCGATGGGAACAAAGAACTGATGAGAACCGCAAGTGTGTTTGTTCACCATAT 360
Qy 362 AATAAAGAAACCACTTCTGAGCCAGACCTGGGCTGTTTACTGTGATGATATTCGAC 421
Db 361 AATAAAGAAACCACTTCTGAGCCAGACCTGGGCTGTTTACTGTGATGATATTCGAC 420
Qy 422 CAAAGCCAAACCCCGGCAAGATACGACGACGACCACTGCGCATGGAATTTCCAGGG 481
Db 421 CAAAGCCAAACCCCGGCAAGATACGACGACGACCACTGCGCATGGAATTTCCAGGG 480
Qy 482 CCGGGATTTTCACTGGCAAAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 541
Db 481 CCGGGATTTTCACTGGCAAAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 540
Qy 542 AACCGCAAGTCTTTGCTCCCTCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 601
Db 541 AACCGCAAGTCTTTGCTCCCTCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 600
Qy 602 AAGCGGAGTGAAGCAGTGTCAAGCATTTTGAAGAAATGCGATAAAGCAAGGTAGAGC 661
Db 601 AAGCGGAGTGAAGCAGTGTCAAGCATTTTGAAGAAATGCGATAAAGCAAGGTAGAGC 660
Qy 662 AATGACCTCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 721
Db 661 AATGACCTCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Qy 722 CAGAATGTGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 781
Db 721 CAGAATGTGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy 782 TCTCACCAGAGATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 841
Db 781 TCTCACCAGAGATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Qy 842 TGTCCAGCTCTCCAGGATGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 901
```

# RESULT 4

AB84485  
ID AAB84485 standard; protein; 414 AA.  
XX AC AAB84485;

05-SEP-2001 (first entry)

Amino acid sequence of FRA16D oxidoreductase (FOR) II gene cDNA.

Cancer associated protein; FOR gene; FRA16D; fragile site; aphidicolin;  
chromosomal rearrangement; cancer; splice variant; DNA instability;  
FRA16D oxidoreductase; neoplasia.

Homo sapiens.

WO200144466-A1.

21-JUN-2001.

15-DEC-2000; 2000WO-AU001539.

16-DEC-1999; 99AU-00004711.

19-APR-2000; 2000AU-00007025.

(WOME-) WOMEN'S & CHILDREN'S HOSPITAL.

Richards R, Ried K, Finnis M, Hobson L, Mangelsdorf M, Dayan S;  
Nancarrow J, Woollatt E, Baker E;

WPI; 2001-398151/42.

N-PSDB; AAH27867.

Novel isolated 16q23.2 nucleic acid molecule, FRA16D oxidoreductase (FOR)  
gene associated with FRA16D site, useful for early diagnosis and  
assessment of risk of cancers associated with the FRA16D region.

Claim 31; Fig 9; 150pp; English.

The present sequence represents a human FRA16D oxidoreductase (FOR) II  
transcript. The FOR gene encodes a cancer associated protein. The FRA16D  
site is a fragile site induced by aphidicolin, which is located within  
the FOR gene. The fragile site is the location of breakpoints of a  
variety of chromosomal rearrangements, and other mutations associated  
with cancers. The FOR protein is expressed as a number of splice  
variants. FOR gene polynucleotide fragments are capable of acting as  
specific primers or probes for detecting cancer associated variations of  
DNA sequence such as a point mutation or small DNA rearrangement

Tue May 25 10:19:07 2004

us-09-

CC associated with the tumour, a breakpoint of one or more chromosomal  
rearrangements associated with the tumour and a pause site within the  
FRA16 gene. FOR nucleic acid molecules are useful as markers to identify  
CC relationship between the fragile site (FRA16D) and the DNA instability in  
CC neoplasia which allows better diagnosis of cancers associated with the  
CC region  
XX  
SQ Sequence 414 AA;

Query Match 99.5%; Score 2191; DB 4; Length 414;  
Best Local Similarity 99.5%; Pred. No. 1.7e-214;  
Matches 412; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	MAALRYAGLDDTDSDELPPGWEERTTKDGGVYVYANHTTEKTQWEHPKTKGKRVAGDLP	60
Db	1	MAALRYAGLDDTDSDELPPGWEERTTKDGGVYVYANHTTEKTQWEHPKTKGKRVAGDLP	60
QY	61	YQWEQETDENGQVFFVDHINKRTTYLDPELAPTVDDNPKPTTRORYDGSSTTAMELQGR	120
Db	61	YQWEQETDENGQVFFVDHINKRTTYLDPELAPTVDDNPKPTTRORYDGSSTTAMELQGR	120
QY	121	DFTGKVVVVTGANSIGIGFETAKSFALHGAHVILACRNMARASAVSRILLEWHKAKVEM	180
Db	121	DFTGKVVVVTGANSIGIGFETAKSFALHGAHVILACRNMARASAVSRILLEWHKAKVEM	180
QY	181	TLDLALLRSVQHHPARAFKAKNVPLHVLVCNAATFALFWSLTKDGLTTQVNHGHFYL	240
Db	181	TLDLALLRSVQHHPARAFKAKNVPLHVLVCNAATFALFWSLTKDGLTTQVNHGHFYL	240
QY	241	QLLODVLCISAPARVIVSSESHRFTDINDSLCKLDFSLSPKNDYMWAMLAYNRSKLCN	300
Db	241	QLLODVLCISAPARVIVSSESHRFTDINDSLCKLDFSLSPKNDYMWAMLAYNRSKLCN	300
QY	301	ILFNSNELHRLSPRGVTSNAVHPGCMWYNIHRSWVYVYLLFTLARPFTKSMQCGAATV	360
Db	301	ILFNSNELHRLSPRGVTSNAVHPGCMWYNIHRSWVYVYLLFTLARPFTKSMQCGAATV	360
QY	361	YCAAVPELEGCGMYFNCCRCMPSPQAQSEETARTLWALSERLIOERLGSQSG	414
Db	361	YCAAVPELEGCGMYFNCCRCMPSPQAQSEETARTLWALSERLIOERLGSQSG	414

## ALIGNMENTS

RESULT 1  
ABG96559  
ID ABG96559 standard; protein; 414 AA.  
AC  
XX ABG96559;  
XX  
DT 12-DEC-2002 (first entry)  
XX  
XX Human short chain dehydrogenase family member WWOX.  
XX  
XX Short chain dehydrogenase; SDR; human; antimicrotoca; pesticide;  
KW herbicide; DHP reductase deficiency; phenylketonuria; galactosaemia III;  
KW dienoyl CoA reductase deficiency; adrenal hyperplasia; ovarian cancer;  
KW adrenogenital syndrome; mineralcorticoid excess syndrome; breast cancer;  
KW male pseudohermaphroditism; Zellweger syndrome; Down's syndrome;  
KW polycystic kidney disease; Alzheimer's disease; retinitis pigmentosa;  
KW retinitis punctata albesens; arterial hypertension; follicular lymphoma;  
KW hepatocarcinogenesis; fungicide; antibiotic.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200212544-A2.  
FN  
XX  
XX 14-FEB-2002.  
PD  
XX  
XX 07-AUG-2001; 2001WO-EP009140.  
PF  
XX  
XX 07-AUG-2000; 2000US-0223436P.  
PR  
XX  
XX (BION-) BIONETWORKS GMEH.  
XX  
XX Wilckens T;  
XX  
XX WPI; 2002-241770/29.  
DR  
XX  
XX Identifying or verifying members of the short chain dehydrogenase (SDR)  
PT family, useful for novel drug development (e.g. for the development of  
PT antimicrotoca, pesticides or herbicides), by employing an algorithm using  
PT core SDR motifs.  
PT  
XX  
XX Disclosure; Fig 4; 168pp; English.  
PS  
XX  
XX The invention relates to identifying or verifying members of the short  
CC chain dehydrogenase (SDR) family comprising employing an algorithm using  
CC core SDR motifs (MT1-MT4 and MV1, MV2 given in the specification) for  
CC searching members of the SDR family. Also included are a member of the  
CC SDR family identified with the method above, a method for providing  
CC modulators for members of the SDR family, a method for evaluation of lead  
CC

Tue May 25 10:19:07 2004

us-09-97

CC -candidates for possible modulators of a member of the SDR family and a  
CC method for detecting clinically relevant polymorphisms or single  
CC nucleotide polymorphisms. The method is useful for screening SDR  
CC sequences and modulators of the SDR family. The method is especially  
CC useful as a platform for novel drug development. The SDRs can serve for  
CC the development of e.g. antimicrotoca, pesticides or herbicides. The  
CC modulators may be especially useful for the prophylaxis, treatment or/and  
CC diagnosis of diseases (e.g. DHP deficiency, phenylketonuria, dienoyl CoA  
CC reductase deficiency, galactosaemia III, adrenal hyperplasia,  
CC adrenogenital syndrome, mineralcorticoid excess syndrome, ovarian cancer,  
CC breast cancer, male pseudohermaphroditism, Zellweger syndrome, polycystic  
CC kidney disease, Alzheimer's disease, retinitis punctata albesens,  
CC retinitis pigmentosa, Down's syndrome, arterial hypertension, follicular  
CC lymphoma and hepatocarcinogenesis) particularly as a fungicide or  
CC antibiotic. The present sequence is one of 39 human SDR family members  
CC identified by the method of the invention  
XX  
SQ Sequence 414 AA;

Query Match	100.0%	Score 2203;	DB 5;	Length 414;
Best Local Similarity	100.0%	Pred. No. 1e-215;		
Matches 414;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	MAALRYAGLDDTDSDELPPGWEERTTKDGMVYVYANHTTEKTQWEHPKTKGRVAGDLP	60		
Db 1	MAALRYAGLDDTDSDELPPGWEERTTKDGMVYVYANHTTEKTQWEHPKTKGRVAGDLP	60		
QY 61	YGEQSTDENGQVFFVDHINKETTYLDPLAPTVDDNPTKPTTORYDGSSTAMEILOQR	120		
Db 61	YGEQSTDENGQVFFVDHINKETTYLDPLAPTVDDNPTKPTTORYDGSSTAMEILOQR	120		
QY 121	DFTGKVVVVTGANSIGIGFETAKSFALHGAHVILACENMARASEAVSRILLEWHKAKVEM	180		
Db 121	DFTGKVVVVTGANSIGIGFETAKSFALHGAHVILACENMARASEAVSRILLEWHKAKVEM	180		
QY 181	TLDLALLRSVQHPAEAFKAKVPLHLVLCNAATFALPWSLTQDGLTTPQVNLGHFVIV	240		
Db 181	TLDLALLRSVQHPAEAFKAKVPLHLVLCNAATFALPWSLTQDGLTTPQVNLGHFVIV	240		
QY 241	QLQDVLCHRSAPARVTVSSSESHRFTDINDSLGKLDPSRLSPKNDYMWALYNSKLCN	300		
Db 241	QLQDVLCHRSAPARVTVSSSESHRFTDINDSLGKLDPSRLSPKNDYMWALYNSKLCN	300		
QY 301	ILFSNELHRLSPRGVTSNAVHPGNMYSNIHRSMMVYTLFLTPARPTKSMQCGAATTV	360		
Db 301	ILFSNELHRLSPRGVTSNAVHPGNMYSNIHRSMMVYTLFLTPARPTKSMQCGAATTV	360		
QY 361	YCAAVPELEGICGMYPNNCCRCMPSPQAQSEHETARTLWALSERLIQERLGSQSG	414		
Db 361	YCAAVPELEGICGMYPNNCCRCMPSPQAQSEHETARTLWALSERLIQERLGSQSG	414		

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 09:56:46 ; Search time 45 Seconds  
(without alignments)  
2902.767 Million cell updates/sec

Title: US-09-978-318B-1

Perfect score: 2203

Sequence: 1 MAALRYAGLDDTSEDELP.....RTLMALSERLQERLGSQSG 414

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPTREMBL 25.\*

2: sp\_archaea.\*

3: sp\_bacteria.\*

4: sp\_fungi.\*

5: sp\_human.\*

6: sp\_invertebrate.\*

7: sp\_mammal.\*

8: sp\_mhc.\*

9: sp\_organelle.\*

10: sp\_plant.\*

11: sp\_rhodent.\*

12: sp\_virus.\*

13: sp\_vertebrate.\*

14: sp\_unclassified.\*

15: sp\_rvirus.\*

16: sp\_bacteriap.\*

17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2203	100.0	414	4 Q9NZC7	Q9nzc7 homo sapien
2	2191	99.5	414	4 Q9NRP5	Q9nrf5 homo sapien
3	2071	94.0	414	11 Q91WL8	Q91wl8 mus musculus
4	2067	93.8	414	11 Q9JLF5	Q9jlf5 mus musculus
5	1869	84.8	363	4 Q9NRF6	Q9nrf6 homo sapien
6	1785	81.0	354	11 Q9D2B3	Q9d2b3 mus musculus
7	1777	80.7	367	11 Q9D339	Q9d339 mus musculus
8	1637.5	74.3	412	13 Q803A8	Q803a8 brachydanio
9	1157	52.5	234	4 Q9BTF8	Q9btf8 homo sapien
10	1149	52.2	242	11 Q920F2	Q920f2 mus musculus
11	1023	46.4	409	5 Q9VL05	Q9vlu5 drosophila
12	942	42.8	191	4 Q9NRK1	Q9nrk1 homo sapien
13	927	42.1	189	4 Q9NPC9	Q9npc9 homo sapien
14	750.5	34.1	213	4 Q96KM3	Q96km3 homo sapien
15	746.5	33.9	311	4 Q96RP2	Q96rf2 homo sapien
16	740	33.6	158	11 Q8C8J6	Q8c8j6 mus musculus

17	541.5	24.6	330	5 Q960C7	Q960c7 drosophila
18	539	24.5	370	5 Q9UAW2	Q9uaw2 caenorhabdi
19	538.5	24.4	329	5 Q19062	Q19062 caenorhabdi
20	538.5	24.4	331	5 Q9V4P9	Q9v4p9 drosophila
21	537.5	24.4	330	5 Q9V4Q3	Q9v4q3 drosophila
22	537	24.4	321	10 Q9LQI8	Q9lqig8 oryza sativ
23	528	24.0	319	5 Q95QN9	Q95qn9 caenorhabdi
24	520.5	23.6	320	10 Q8L9T6	Q8l9t6 arabidopsis
25	520.5	23.6	325	5 Q8MZG9	Q8mzg9 drosophila
26	514.5	23.4	316	10 Q8H1Q6	Q8h1q6 arabidopsis
27	510.5	23.2	300	5 Q9V4Q2	Q9v4q2 drosophila
28	509.5	23.1	296	5 Q8MKN1	Q8mkn1 drosophila
29	506.5	23.0	293	11 Q9R1R9	Q9r1r9 mus musculu
30	506.5	23.0	300	11 Q9R1R8	Q9r1r8 mus musculu
31	497	22.6	317	10 Q9LDY7	Q9ldy7 arabidopsis
32	497	22.6	321	10 Q80924	Q80924 arabidopsis
33	495	22.5	331	10 Q8RWJ2	Q8rwj2 arabidopsis
34	493	22.4	319	10 Q94LU0	Q94lu0 oryza sativ
35	493	22.4	319	10 Q7XCH3	Q7xch3 oryza sativ
36	490.5	22.3	370	4 Q12953	Q12953 homo sapien
37	489	22.2	331	10 Q8LCE7	Q8lce7 arabidopsis
38	485.5	22.0	309	13 Q7T3A8	Q7t3a8 brachydanio
39	485.5	22.0	406	5 Q9W404	Q9w404 drosophila
40	478	21.7	379	5 O45680	O45680 caenorhabdi
41	476.5	21.6	350	10 Q9LZ48	Q9lzl48 arabidopsis
42	472	21.4	336	16 Q9RR99	Q9rr99 deinococcus
43	469	21.3	314	10 Q7XMT5	Q7xmt5 oryza sativ
44	469	21.3	320	10 P81259	P81259 brassica na
45	447.5	20.3	294	10 O81739	O81739 arabidopsis

#### ALIGNMENTS

#### RESULT 1

ID	Q9NZC7	PRELIMINARY;	PRT;	414 AA.
AC	Q9NZC7;			
DC	01-OCT-2000 (TREMELrel. 15, Created)			
DT	01-OCT-2000 (TREMELrel. 15, Last sequence update)			
DT	01-OCT-2003 (TREMELrel. 25, Last annotation update)			
DE	WW domain-containing protein WMOX (WW domain-containing oxidoreductase isoform FORII).			
GN	WMOX.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20246348; PubMed=10786676;			
RA	Bednarek A.K., Laflin K.J., Daniel R.L., Liao Q., Hawkins K.A.,			
RA	Aldaz C.M.;			
RT	"WMOX, a novel WW domain-containing protein mapping to human			
RT	chromosome 16q23.3-24.1, a region frequently affected in breast			
RT	cancer.";			
RL	Cancer Res. 60:2140-2145(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	PubMed=11572989;			
RA	Paige A.J.W., Taylor K.J., Taylor C., Hallier S.G., Farrington S.,			
RA	Scott D., Porteous D.J., Smyth J.F., Gabra H., Watson J.E.V.;			
RT	"WMOX: A candidate tumor suppressor gene involved in multiple tumor			
RT	types.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 98:11417-11422(2001).			
CC	-!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES			
CC	(SDR) FAMILY.			
DR	EMBL; AF211943; AAF27049.1; -			
DR	EMBL; AF325432; AAL05449.1; -			
DR	EMBL; AF325423; AAL05449.1; JOINED.			
DR	EMBL; AF325424; AAL05449.1; JOINED.			
DR	EMBL; AF325425; AAL05449.1; JOINED.			
DR	EMBL; AF325426; AAL05449.1; JOINED.			



```
DR EMBL; AF325427; AAL05449.1; JOINED.
DR EMBL; AF325428; AAL05449.1; JOINED.
DR EMBL; AF325430; AAL05449.1; JOINED.
DR EMBL; AF325431; AAL05449.1; JOINED.
DR HSSP; Q13526; 1PIN.
DR GO; GO:0005489; F:electron transporter activity; TAS.
DR GO; GO:0005515; F:protein binding; TAS.
DR GO; GO:0008202; P:steroid metabolism; TAS.
DR InterPro; IPR002198; ADH short.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00106; adh_short; 1.
DR Pfam; PF00397; WW_2.
DR SMART; SM00456; WW; 2.
DR PROSITE; PS01159; WW_DOMAIN_1; 2.
DR PROSITE; PS0020; WW_DOMAIN_2; 2.
KW Oxidoreductase.
SQ SEQUENCE 414 AA; 46676 MW; E4D9A649E6CE05DF CRC64;

Query Match 100.0%; Score 2203; DB 4; Length 414;
Best Local Similarity 100.0%; Pred. No. 1.5e-176;
Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAALRYAGLDGDDTSEDELPPEGWEERTTKDGVVYVYANHTEKTOWEHPKTKRKRKRVAGDLP 60
DB 1 MAALRYAGLDGDDTSEDELPPEGWEERTTKDGVVYVYANHTEKTOWEHPKTKRKRKRVAGDLP 60

QY 61 YGWEQETDENGQVFFVDHINKRTTYLDPRLAFTVDNDNPTKPTTRQRYDGSSTTAMEILOQR 120
DB 61 YGWEQETDENGQVFFVDHINKRTTYLDPRLAFTVDNDNPTKPTTRQRYDGSSTTAMEILOQR 120

QY 121 DFTGKVVVVTGANGSIGGFETAKSFALHGAHVILACRNMARASEAVSRILEEWHKAKVEAM 180
DB 121 DFTGKVVVVTGANGSIGGFETAKSFALHGAHVILACRNMARASEAVSRILEEWHKAKVEAM 180

QY 181 TLDLALLRSVQHFAFAFAKAKVPLHVLVCNAATFALPWSLTQDGLTTTQVNHHLGHFYLV 240
DB 181 TLDLALLRSVQHFAFAFAKAKVPLHVLVCNAATFALPWSLTQDGLTTTQVNHHLGHFYLV 240

QY 241 QLLQDVLCRSAPARVIVVSSSEHRTDINDSLGKLDIFSRLSPKNDYMWALAYNRSKLCN 300
DB 241 QLLQDVLCRSAPARVIVVSSSEHRTDINDSLGKLDIFSRLSPKNDYMWALAYNRSKLCN 300

QY 301 ILFNSLHRLSPRGVTSNAVHPGNMYSNIHRSWWVYLLFTLAPPTKSMQOQAATTV 360
DB 301 ILFNSLHRLSPRGVTSNAVHPGNMYSNIHRSWWVYLLFTLAPPTKSMQOQAATTV 360

QY 361 YCAAVPELEGLGMYFNCCRCMPSEAOSEETARTLWALSERLIQERLGSQSG 414
DB 361 YCAAVPELEGLGMYFNCCRCMPSEAOSEETARTLWALSERLIQERLGSQSG 414

RESULT 2
Q9NRP5 Q9NRP5 PRELIMINARY; PRT; 414 AA.
AC Q9NRP5;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE For II protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20320695; PubMed=10861292;
RA Ried K., Finniss M., Hobson L., Mangelndorf M., Dayan S.,
RA Nancarrow J.K., Woolatt E., Kremmidiotis G., Gardner A., Venter D.,
RA Baker E., Richards R.I.;
RT "Common chromosomal fragile site FRA16D sequence: Identification of the
RT FOR gene spanning FRA16D and homozygous deletions and translocation
RT breakpoints in cancer cells.";
RL Hum. Mol. Genet. 9:1651-1663 (2000).
```

```
CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC EMBL; AF227527; AAF82054.1; -.
CC HSSP; Q13526; 1PIN.
CC Genew; HGNC:12799; WWOX.
CC GO; GO:0016491; F:oxidoreductase activity; IEA.
CC GO; GO:0008152; P:metabolism; IEA.
CC InterPro; IPR002198; ADH short.
CC InterPro; IPR001202; WW_Rsp5_WWP.
CC Pfam; PF00106; adh_short; 1.
CC Pfam; PF00397; WW_2.
CC SMART; SM00456; WW; 2.
CC PROSITE; PS01159; WW_DOMAIN_1; 2.
CC PROSITE; PS0020; WW_DOMAIN_2; 2.
KW Oxidoreductase.
SQ SEQUENCE 414 AA; 46680 MW; 04D9B358F87F0E75 CRC64;

Query Match 99.5%; Score 2191; DB 4; Length 414;
Best Local Similarity 99.5%; Pred. No. 1.6e-175;
Matches 412; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAALRYAGLDGDDTSEDELPPEGWEERTTKDGVVYVYANHTEKTOWEHPKTKRKRKRVAGDLP 60
DB 1 MAALRYAGLDGDDTSEDELPPEGWEERTTKDGVVYVYANHTEKTOWEHPKTKRKRKRVAGDLP 60

QY 61 YGWEQETDENGQVFFVDHINKRTTYLDPRLAFTVDNDNPTKPTTRQRYDGSSTTAMEILOQR 120
DB 61 YGWEQETDENGQVFFVDHINKRTTYLDPRLAFTVDNDNPTKPTTRQRYDGSSTTAMEILOQR 120

QY 121 DFTGKVVVVTGANGSIGGFETAKSFALHGAHVILACRNMARASEAVSRILEEWHKAKVEAM 180
DB 121 DFTGKVVVVTGANGSIGGFETAKSFALHGAHVILACRNMARASEAVSRILEEWHKAKVEAM 180

QY 181 TLDLALLRSVQHFAFAFAKAKVPLHVLVCNAATFALPWSLTQDGLTTTQVNHHLGHFYLV 240
DB 181 TLDLALLRSVQHFAFAFAKAKVPLHVLVCNAATFALPWSLTQDGLTTTQVNHHLGHFYLV 240

QY 241 QLLQDVLCRSAPARVIVVSSSEHRTDINDSLGKLDIFSRLSPKNDYMWALAYNRSKLCN 300
DB 241 QLLQDVLCRSAPARVIVVSSSEHRTDINDSLGKLDIFSRLSPKNDYMWALAYNRSKLCN 300

QY 301 ILFNSLHRLSPRGVTSNAVHPGNMYSNIHRSWWVYLLFTLAPPTKSMQOQAATTV 360
DB 301 ILFNSLHRLSPRGVTSNAVHPGNMYSNIHRSWWVYLLFTLAPPTKSMQOQAATTV 360

QY 361 YCAAVPELEGLGMYFNCCRCMPSEAOSEETARTLWALSERLIQERLGSQSG 414
DB 361 YCAAVPELEGLGMYFNCCRCMPSEAOSEETARTLWALSERLIQERLGSQSG 414

RESULT 3
Q91WL8 Q91WL8 PRELIMINARY; PRT; 414 AA.
AC Q91WL8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE WW domain-containing oxidoreductase.
OS WWOX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mullerian duct includes surrounding region;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
```

RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RL 60,770 full-length cDNAs."; J. Biol. Chem. 276:3361-3370(2001).  
 CC Nature 420:563-573(2002).  
 CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
 (SDR) FAMILY.  
 DR EMBL; AF187014; AAF31693.1; -.  
 DR HSSP; Q13526; 1PIN.  
 DR MGD; MGI:1931237; Wwox.  
 DR GO; GO:0005739; C:mitochondrion; IDA.  
 DR GO; GO:0005634; C:nucleus; IDA.  
 DR GO; GO:0006917; P:induction of apoptosis; IDA.  
 DR InterPro; IPR002198; ADH short.  
 DR InterPro; IPR001202; WW\_Rsp5\_WWP.  
 DR Pfam; PF00106; adh short; 1.  
 DR Pfam; PF00397; WW\_2.  
 DR SMART; SM00456; WW; 2.  
 DR PROSITE; PS01159; WW DOMAIN 1; 2.  
 DR PROSITE; PS00020; WW DOMAIN 2; 2.  
 KW Oxidoreductase.  
 SQ SEQUENCE 414 AA; 46512 MW; 3C83AE3085B6A931 CRC64;

Query Match 94.0%; Score 2071; DB 11; Length 414;  
 Best Local Similarity 93.9%; Pred. No. 1.9e-165;  
 Matches 388; Conservative 10; Mismatches 15; Indels 0; Gaps 0;  
 QY 1 MAALRYAGLDGTDSEDELPDGGWEERTTKDGVVYANHTTEKTQWHPKTKGKRVAGDLP 60  
 DB 1 MAALRYAGLDGTDSEDELPDGGWEERTTKDGVVYANHTTEKTQWHPKTKGKRVAGDLP 60  
 QY 61 YGWEQETDENGQVFPVDHINKRTTYLDPRLAFTVDNDNPKPTTRQRYDGSSTTAMEILQGR 120  
 DB 61 YGWEQETDENGQVFPVDHINKRTTYLDPRLAFTVDNDNPKPTTRQRYDGSSTTAMEILQGR 120  
 QY 121 DFTGKVVVVTGANGSIGIGPETAKSPALHGAHVILACRNMRASEAVSRILEEWHKAKVEAM 180  
 DB 121 DFTGKVVVVTGANGSIGIGPETAKSPALHGAHVILACRNMRASEAVSRILEEWHKAKVEAM 180  
 QY 181 TLDLALLRSVOHEAFKAKVPLHVLVCNAATFALPMSLTQDGLTTFQVNHGHHFYLV 240  
 DB 181 TLDLALLRSVOHEAFKAKVPLHVLVCNAATFALPMSLTQDGLTTFQVNHGHHFYLV 240  
 QY 241 QLLQDVLCRSPARVIVVSSSHRFTDINDSLGKIDFSRLSPTKNDYWLAYNRSKLCN 300  
 DB 241 QLLQDVLCRSPARVIVVSSSHRFTDINDSLGKIDFSRLSPTKNDYWLAYNRSKLCN 300  
 QY 301 ILFSNELHRLSPRGVTSNAVHPGNMYSNTHRSWVYVTLFTLARPTKSMQQAATTV 360  
 DB 301 ILFSNELHRLSPRGVTSNAVHPGNMYSNTHRSWVYVTLFTLARPTKSMQQAATTV 360  
 QY 361 YCAAVPELEGGLGMYFNCCRCMPSPQAQSEETARTLWALSERLIOERLGSQS 413  
 DB 361 YCAVAPLEGGLGMYFNCCRCCLPSEEAQSEETARALWELSERLIOERLGSQS 413

RESULT 4  
 Q9JLF5 PRELIMINARY; PRT; 414 AA.  
 AC Q9JLF5;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE WW-domain oxidoreductase.  
 GN WWOX.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21264809; PubMed=11058590;  
 RA Chang N.-S., Pratt N., Heath J., Schultz L., Sleeve D., Carey G.B.,  
 RA Zevotek N.;

RT "Hyaluronidase Induction of a WW Domain-Containing Oxidoreductase that  
 RT Enhances Tumor Necrosis Factor Cytotoxicity."; J. Biol. Chem. 276:3361-3370(2001).  
 CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
 (SDR) FAMILY.  
 DR EMBL; AF187014; AAF31693.1; -.  
 DR HSSP; Q13526; 1PIN.  
 DR MGD; MGI:1931237; Wwox.  
 DR GO; GO:0005739; C:mitochondrion; IDA.  
 DR GO; GO:0005634; C:nucleus; IDA.  
 DR GO; GO:0006917; P:induction of apoptosis; IDA.  
 DR InterPro; IPR002198; ADH short.  
 DR InterPro; IPR001202; WW\_Rsp5\_WWP.  
 DR Pfam; PF00106; adh short; 1.  
 DR Pfam; PF00397; WW\_2.  
 DR SMART; SM00456; WW; 2.  
 DR PROSITE; PS01159; WW DOMAIN 1; 2.  
 DR PROSITE; PS00020; WW DOMAIN 2; 2.  
 KW Oxidoreductase.  
 SQ SEQUENCE 414 AA; 46484 MW; 1B8BA12882B6BE3P CRC64;  
 Query Match 93.8%; Score 2067; DB 11; Length 414;  
 Best Local Similarity 93.7%; Pred. No. 4.1e-165;  
 Matches 387; Conservative 10; Mismatches 16; Indels 0; Gaps 0;  
 QY 1 MAALRYAGLDGTDSEDELPDGGWEERTTKDGVVYANHTTEKTQWHPKTKGKRVAGDLP 60  
 DB 1 MAALRYAGLDGTDSEDELPDGGWEERTTKDGVVYANHTTEKTQWHPKTKGKRVAGDLP 60  
 QY 61 YGWEQETDENGQVFPVDHINKRTTYLDPRLAFTVDNDNPKPTTRQRYDGSSTTAMEILQGR 120  
 DB 61 YGWEQETDENGQVFPVDHINKRTTYLDPRLAFTVDNDNPKPTTRQRYDGSSTTAMEILQGR 120  
 QY 121 DFTGKVVVVTGANGSIGIGPETAKSPALHGAHVILACRNMRASEAVSRILEEWHKAKVEAM 180  
 DB 121 DFTGKVVVVTGANGSIGIGPETAKSPALHGAHVILACRNMRASEAVSRILEEWHKAKVEAM 180  
 QY 181 TLDLALLRSVOHEAFKAKVPLHVLVCNAATFALPMSLTQDGLTTFQVNHGHHFYLV 240  
 DB 181 TLDLALLRSVOHEAFKAKVPLHVLVCNAATFALPMSLTQDGLTTFQVNHGHHFYLV 240  
 QY 241 QLLQDVLCRSPARVIVVSSSHRFTDINDSLGKIDFSRLSPTKNDYWLAYNRSKLCN 300  
 DB 241 QLLQDVLCRSPARVIVVSSSHRFTDINDSLGKIDFSRLSPTKNDYWLAYNRSKLCN 300  
 QY 301 ILFSNELHRLSPRGVTSNAVHPGNMYSNTHRSWVYVTLFTLARPTKSMQQAATTV 360  
 DB 301 ILFSNELHRLSPRGVTSNAVHPGNMYSNTHRSWVYVTLFTLARPTKSMQQAATTV 360  
 QY 361 YCAAVPELEGGLGMYFNCCRCMPSPQAQSEETARTLWALSERLIOERLGSQS 413  
 DB 361 YCAVAPLEGGLGMYFNCCRCCLPSEEAQSEETARALWELSERLIOERLGSQS 413

RESULT 5  
 Q9NRF6 PRELIMINARY; PRT; 363 AA.  
 AC Q9NRF6;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE For I protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20320695; PubMed=10861292;  
 RA Ried K., Finniss M., Hobson L., Mangelsdorf M., Dayan S.,  
 RA Nancarrow J.K., Woolatt E., Kremmidiotis G., Gardner A., Venter D.,  
 RA Baker E., Richards R.I.;  
 RT "Common chromosomal fragile site FRA16D sequence: Identification of the

RT FOR gene spanning FRA16D and homozygous deletions and translocation  
RT breakpoints in cancer cells.;

RL Hum. Mol. Genet. 9:1651-1663(2000).

CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
(SDR) FAMILY.

DR EMBL; AF227526; AAF82053.1; -.

DR HSSP; Q13526; 1PIN.

DR GO; GO:0016491; F:oxidoreductase activity; IEA.

DR GO; GO:0008152; P:metabolism; IEA.

DR InterPro; IPR002198; ADH short.

DR InterPro; IPR001202; WW\_Rep5\_WWP.

DR Pfam; PF00106; adh\_short; 1.

DR Pfam; PF00397; WW; 2.

DR SMART; SM00456; WW; 2.

DR PROSITE; PS01159; WW\_DOMAIN\_1; 2.

DR PROSITE; PS00020; WW\_DOMAIN\_2; 2.

KW Oxidoreductase.

SQ SEQUENCE 363 AA; 41157 MW; C8DE29B7F11781D6 CRC64;

Query Match 84.8%; Score 1869; DB 4; Length 363;

Best Local Similarity 99.7%; Pred. No. 1.5e-148;

Matches 351; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAALRYAGLDDTSDDELPPGWEERTTKDGWVYANHTTEKTQWEHPKTKGKRVRVAGDLP 60

DB 1 MAALRYAGLDDTSDDELPPGWEERTTKDGWVYANHTTEKTQWEHPKTKGKRVRVAGDLP 60

QY 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDNPKPTTRQYDGTSTTAMELQOR 120

DB 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDNPKPTTRQYDGTSTTAMELQOR 120

QY 121 DFTGKVVVVTGANGSIGFETAKSFALHGAHVILACRNMARASEAVSRILEEWHKAKVEAM 180

DB 121 DFTGKVVVVTGANGSIGFETAKSFALHGAHVILACRNMARASEAVSRILEEWHKAKVEAM 180

QY 181 TLDLALLRSVQHFAEAFKAKNVLPLVLCNAATFALPWSLTKDGLTTQVNHGHFYLV 240

DB 181 TLDLALLRSVQHFAEAFKAKNVLPLVLCNAATFALPWSLTKDGLTTQVNHGHFYLV 240

QY 241 QLLQDVLCRSAPRVIVVSSSHRFTDINDSLGKLDLRLSTKNDYMWAMLYNRSKLCN 300

DB 241 QLLQDVLCRSAPRVIVVSSSHRFTDINDSLGKLDLRLSTKNDYMWAMLYNRSKLCN 300

QY 301 ILFSNELHRLSPRGVTSNAVHPGNMYSNIHRSWMVYTLFLTPRPFTKSM 352

DB 301 ILFSNELHRLSPRGVTSNAVHPGNMYSNIHRSWMVYTLFLTPRPFTKSM 352

RESULT 6

Q9D2B3

ID Q9D2B3 PRELIMINARY; PRT; 354 AA.

AC Q9D2B3;

DT 01-JUN-2001 (TREMELrel. 17, Created)

DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)

DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)

DE 5330426P09Rik protein.

GN WWOX OR 5330426P09Rik.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Pituitary;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

RA Saito T., Okazaki Y., Gofjohori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Giasi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wyszaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
RA Hayashizaki Y.,  
RT "Functional annotation of a full-length mouse cDNA collection.;"  
RL Nature 409:685-690(2001).

CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
(SDR) FAMILY.

DR EMBL; AK019911; BAB31911.1; -.

DR HSSP; Q13526; 1PIN.

DR MGD; MGI:1931237; Wwox.

DR GO; GO:0016491; F:oxidoreductase activity; IEA.

DR GO; GO:0008152; P:metabolism; IEA.

DR InterPro; IPR002198; ADH short.

DR InterPro; IPR001202; WW\_Rep5\_WWP.

DR Pfam; PF00106; adh\_short; 1.

DR Pfam; PF00397; WW; 2.

DR SMART; SM00456; WW; 2.

DR PROSITE; PS01159; WW\_DOMAIN\_1; 2.

DR PROSITE; PS00020; WW\_DOMAIN\_2; 2.

KW Oxidoreductase.

SQ SEQUENCE 354 AA; 40018 MW; 41CBA77635E5E985 CRC64;

Query Match 81.0%; Score 1785; DB 11; Length 354;

Best Local Similarity 95.2%; Pred. No. 1.6e-141;

Matches 335; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 1 MAALRYAGLDDTSDDELPPGWEERTTKDGWVYANHTTEKTQWEHPKTKGKRVRVAGDLP 60

DB 1 MAALRYAGLDDTSDDELPPGWEERTTKDGWVYANHTTEKTQWEHPKTKGKRVRVAGDLP 60

QY 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDNPKPTTRQYDGTSTTAMELQOR 120

DB 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDNPKPTTRQYDGTSTTAMELQOR 120

QY 121 DFTGKVVVVTGANGSIGFETAKSFALHGAHVILACRNMARASEAVSRILEEWHKAKVEAM 180

DB 121 DFTGKVVVVTGANGSIGFETAKSFALHGAHVILACRNMARASEAVSRILEEWHKAKVEAM 180

QY 181 TLDLALLRSVQHFAEAFKAKNVLPLVLCNAATFALPWSLTKDGLTTQVNHGHFYLV 240

DB 181 TLDLALLRSVQHFAEAFKAKNVLPLVLCNAATFALPWSLTKDGLTTQVNHGHFYLV 240

QY 241 QLLQDVLCRSAPRVIVVSSSHRFTDINDSLGKLDLRLSTKNDYMWAMLYNRSKLCN 300

DB 241 QLLQDVLCRSAPRVIVVSSSHRFTDINDSLGKLDLRLSTKNDYMWAMLYNRSKLCN 300

QY 301 ILFSNELHRLSPRGVTSNAVHPGNMYSNIHRSWMVYTLFLTPRPFTKSM 352

DB 301 ILFSNELHRLSPRGVTSNAVHPGNMYSNIHRSWMVYTLFLTPRPFTKSM 352

RESULT 7

Q9D339

ID Q9D339 PRELIMINARY; PRT; 367 AA.

AC Q9D339;

DT 01-JUN-2001 (TREMELrel. 17, Created)

DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)

DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)

DE 9030416C10Rik protein.

GN WWOX OR 9030416C10Rik.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CS7BL/6J; TISSUE=Colon;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kwai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Marzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,  
 RA Suzuki H., Iyoy-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohatsu S.,  
 RA Hayashizaki Y.;  
 RA "Functional annotation of a full-length mouse cDNA collection.";  
 RT Nature 409:685-690(2001).  
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
 CC (SDR) FAMILY.  
 DR EMBL; AK018507; BAB31244.1; -.  
 DR HSSP; Q13526; 1PIN.  
 DR MGD; MGI:1931237; Wwox.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR002198; ADH\_short.  
 DR InterPro; IPR001202; WW\_Reps\_WWP.  
 DR Pfam; PF00106; adh\_short; 1.  
 DR Pfam; PF00397; WW; 2.  
 DR SMART; SM00456; WW; 2.  
 DR PROSITE; PS01159; WW\_DOMAIN\_1; 1.  
 DR PROSITE; PS0020; WW\_DOMAIN\_2; 2.  
 DR OXIDOREDUCTASE.  
 KW OXIDOREDUCTASE.  
 SQ SEQUENCE 367 AA; 41499 MW; B29A368E793B2C0F CRC64;

Query Match 80.7%; Score 1777; DB 11; Length 367;  
 Best Local Similarity 94.6%; Pred. No. 7.9e-141; Indels 0; Gaps 0;  
 Matches 334; Conservative 9; Mismatches 10;

Qy 1 MAALRYAGLDDTSEDELPDPCWEERTTKDGVVYVYANHTTEKTQWEHPTGKRKRVAGDLP 60  
 Db 1 MAALRYAGLDDTSEDELPDPCWEERTTKDGVVYVYANHTTEKTQWEHPTGKRKRVAGDLP 60  
 Qy 61 YGWEQETDENGQVFVVDHINKRTYLDPLAFTVDDNPTKPTTQRQYDGSSTAMEILQGR 120  
 Db 61 YGWEQETDENGQVFVVDHINKRTYLDPLAFTVDDNPTKPTTQRQYDGSSTAMEILQGR 120  
 Qy 121 DFTGKVVVVVTGANGSIGFETAKSPALHGAHVILACRNARASEAVSRILEWHKAKVEAM 180  
 Db 121 DFTGKVVVVVTGANGSIGFETAKSPALHGAHVILACRNARASEAVSRILEWHKAKVEAM 180  
 Qy 181 TLDLALRSVQHFAEAFKAKNVPLHVCNAAATFALPWSLTGKLETTFQVNHILGHFYL 240  
 Db 181 TLDLALRSVQHFAEAFKAKNVPLHVCNAAATFALPWSLTGKLETTFQVNHILGHFYL 240  
 Qy 241 QLQDVLCRSAPARVIVVSSSHRFTDINDSLGKLDPSRLSPTKNDYVWMLAYNRSLCN 300  
 Db 241 QLQDVLCRSAPARVIVVSSSHRFTDINDSLGKLDPSRLSPTKNDYVWMLAYNRSLCN 300  
 Qy 301 ILFSNELHRLSPRGVTSNAVHPGNMYSNTHRSWVYVTLFTLARPTKSMQ 353  
 Db 301 ILFSNELHRLSPRGVTSNAVHPGNMYSNTHRSWVYVTLFTLARPTKSMQ 353

RESULT 8  
 Q803A8 PRELIMINARY; PRT; 412 AA.  
 AC Q803A8;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Similar to WW domain containing oxidoreductase.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC044560; AAH44560.1; -.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR002198; ADH\_short.  
 DR InterPro; IPR001202; WW\_Reps\_WWP.  
 DR Pfam; PF00106; adh\_short; 1.  
 DR Pfam; PF00397; WW; 2.  
 DR SMART; SM00456; WW; 2.  
 DR PROSITE; PS01159; WW\_DOMAIN\_1; 1.  
 DR PROSITE; PS0020; WW\_DOMAIN\_2; 2.  
 DR OXIDOREDUCTASE.  
 KW OXIDOREDUCTASE.  
 SQ SEQUENCE 412 AA; 46322 MW; D4D10CA809138B8D CRC64;

Query Match 74.3%; Score 1637.5; DB 13; Length 412;  
 Best Local Similarity 72.8%; Pred. No. 4.9e-129;  
 Matches 297; Conservative 51; Mismatches 57; Indels 3; Gaps 1;

Qy 1 MAALRYAGLDDTSEDELPDPCWEERTTKDGVVYVYANHTTEKTQWEHPTGKRKRVAGDLP 60  
 Db 1 MAALRYAGLDDTSEDELPDPCWEERTTKDGVVYVYANHTTEKTQWEHPTGKRKRVAGDLP 60  
 Qy 61 YGWEQETDENGQVFVVDHINKRTYLDPLAFTVDDNPTKPTTQRQYDGSSTAMEILQGR 120  
 Db 61 YGWEQETDENGQVFVVDHINKRTYLDPLAFTVDDNPTKPTTQRQYDGSSTAMEILQGR 120  
 Qy 121 DFTGKVVVVVTGANGSIGFETAKSPALHGAHVILACRNARASEAVSRILEWHKAKVEAM 180  
 Db 121 DFTGKVVVVVTGANGSIGFETAKSPALHGAHVILACRNARASEAVSRILEWHKAKVEAM 180  
 Qy 181 TLDLALRSVQHFAEAFKAKNVPLHVCNAAATFALPWSLTGKLETTFQVNHILGHFYL 240  
 Db 181 TLDLALRSVQHFAEAFKAKNVPLHVCNAAATFALPWSLTGKLETTFQVNHILGHFYL 240  
 Qy 241 QLQDVLCRSAPARVIVVSSSHRFTDINDSLGKLDPSRLSPTKNDYVWMLAYNRSLCN 300  
 Db 241 QLQDVLCRSAPARVIVVSSSHRFTDINDSLGKLDPSRLSPTKNDYVWMLAYNRSLCN 300  
 Qy 301 ILFSNELHRLSPRGVTSNAVHPGNMYSNTHRSWVYVTLFTLARPTKSMQ 360  
 Db 301 ILFSNELHRLSPRGVTSNAVHPGNMYSNTHRSWVYVTLFTLARPTKSMQ 360  
 Qy 361 YCAVPELEGIGGMVFNCCRCMPSPEAQSEETARTLWALSERLIQER 408  
 Db 361 YCAVPELEGIGGMVFNCCRCMPSPEAQSEETARTLWALSERLIQER 408  
 Qy 358 YCAVPELEGIGGMVFNCCRCMPSPEAQSEETARTLWALSERLIQER 405  
 Db 358 YCAVPELEGIGGMVFNCCRCMPSPEAQSEETARTLWALSERLIQER 405

RESULT 9  
 Q9BT8 PRELIMINARY; PRT; 234 AA.  
 AC Q9BT8;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Similar to WW domain-containing oxidoreductase (Wwoxdelta6-8).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Artiodactyla; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;

Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 [2]  
 SEQUENCE FROM N.A.  
 RA Bednarek A.K., Keck-Waggoner C.L., Daniel R.L., Lafin K.J.,  
 RA Kiguchi K., Brenner A.J., Aldaz C.M.;  
 RT "WMOX, the FRA16D gene, behaves as a suppressor of tumor growth."  
 RN Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 [3]  
 SEQUENCE FROM N.A.  
 RA Kalline N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,  
 RA Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,  
 RA Phelan M., Farmer A.;  
 RT "Cloning of human full-length cDNAs in BD Creator (TM) System Donor  
 vector."  
 RN Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
 (SDR) FAMILY  
 DR EMBL; BC003184; AA03184.1; -;  
 DR EMBL; AF395123; AAK81727.1; -;  
 DR EMBL; BT007445; AAK36113.1; -;  
 DR HSSP; Q13526; 1PIN.  
 DR GO; GO:0016491; P:oxidoreductase activity; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR002198; ADH\_short.  
 DR Pfam; PF00106; adh\_short; WW.  
 DR Pfam; PF00397; WW; 2.  
 DR SMART; SM00456; WW; 2.  
 DR PROSITE; PS01159; WW\_DOMAIN 1; 2.  
 DR PROSITE; PS00020; WW\_DOMAIN 2; 2.  
 KW Oxidoreductase.  
 SQ SEQUENCE 234 AA; 26148 MW; C69FB63E87635F4 CRC64;  
 Query Match 52.5%; Score 1157; DB 4; Length 234;  
 Best Local Similarity 56.5%; Pred. No. 5e-89;  
 Matches 234; Conservative 0; Mismatches 0; Indels 180; Gaps 1;  
 QY 1 MAALRYAGLDLTDSEDELPPGWEERTTKDGWVYVYANHTTEKTQWEHPKTKGKRVAGDLP 60  
 Db 1 MAALRYAGLDLTDSEDELPPGWEERTTKDGWVYVYANHTTEKTQWEHPKTKGKRVAGDLP 60  
 QY 61 YWQETDENGQVFFVDHINKETTYLDPLATVDNDNPKPTTRDYDGTSTAMEILQGR 120  
 Db 61 YWQETDENGQVFFVDHINKETTYLDPLATVDNDNPKPTTRDYDGTSTAMEILQGR 120  
 QY 121 DFTGKVVVVTGANSIGFTAKSEALHGAHVILACRNMARASEAVSRILEEWHKAKVEAM 180  
 Db 121 DFTGKVVVVTGANSIGFTAKSEALHGAHVILACRNMARASEAVSRILEEWHKAKVEAM 172  
 QY 181 TLDLALLRSVORFAEAFKAKNVPLHLVCNAATFALPWSLTQDGLTTPQVNLGHFYLIV 240  
 Db 173 ----- 172  
 QY 241 QLLQDVLCESAPARVIVVSSSHRFTDINDSLGKDFSLPTKNDYWMALYNRSKLCN 300  
 Db 173 ----- 172  
 QY 301 ILFSLNELHRLSPRGVTSNAVHPGNMYSNIHRSWWVYTLFLARPFTKSMQOGAATTV 360  
 Db 173 ----- QOGAATTV 180  
 QY 361 YCAAVELEGLOGMYFNCCRCMPSEPAQSEETARTLWALSERLIQERLGSQSG 414  
 Db 181 YCAAVELEGLOGMYFNCCRCMPSEPAQSEETARTLWALSERLIQERLGSQSG 234

RESULT 10

Q920Y2

ID Q920Y2 PRELIMINARY; PRT; 242 AA.

AC Q920Y2;

DT 01-DEC-2001 (T-EMBLrel. 19, Created)

DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)

DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)

DE WW-domain oxidoreductase (Fragment).  
 GN WWOX.OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]

RP SEQUENCE FROM N.A.

RA Krummel K.A., Denison S.R., Calhoun E., Smith D.I.;  
 RT "The Common Fragile Site FRA16D and its Associated Gene WWOX are  
 RT Highly Conserved in Mouse at Fra8E1.";  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY046556; AAL03972.1; -;

DR EMBL; AY046553; AAL03972.1; JOINED.

DR EMBL; AY046554; AAL03972.1; JOINED.

DR EMBL; AY046555; AAL03972.1; JOINED.

DR MGD; MGI:1931237; Wwox.

DR GO; GO:0005739; C:mitochondrion; IEA.

DR GO; GO:0005634; C:nucleus; IEA.

DR GO; GO:0006917; P:induction of apoptosis; IEA.

FT NON TER 1

SQ SEQUENCE 242 AA; 27036 MW; 35F91C09F8D815DA CRC64;

Query Match 52.2%; Score 1149; DB 11; Length 242;  
 Best Local Similarity 90.5%; Pred. No. 2.5e-88;  
 Matches 218; Conservative 7; Mismatches 16; Indels 0; Gaps 0;QY 173 HKAKVEAMTLDLALLRSVQHFABEAPKAKNVPLHLVCNAATFALPWSLTQDGLTTPQV 232  
 Db 1 HKAKVEAMTLDLALLRSVQHFABEAPKAKNVPLHLVCNAATFALPWSLTQDGLTTPQV 60

QY 233 HIGHFYLVQLLQDVLCRSAPARVIVVSSSHRFTDINDSLGKDFSLPTKNDYWMALIA 292

Db 61 HIGHFYLVQLLQDVLCRSAPARVIVVSSSHRFTDINDSLGKDFSLPTKNDYWMALIA 120

QY 293 YNRSKLCNLFSLNELHRLSPRGVTSNAVHPGNMYSNIHRSWWVYTLFLARPFTKSM 352

Db 121 YNRSKLCNLFSLNELHRLSPRGVTSNAVHPGNMYSNIHRSWWVYTLFLARPFTKSM 180

QY 353 OOGAATTVCYCAVPELEGLOGMYFNCCRCMPSEPAQSEETARTLWALSERLIQERLGSQ 412

Db 181 OOGAATTVCYCAVPELEGLOGMYFNCCRCMPSEPAQSEETARTLWALSERLIQERLGSQ 240

QY 413 S 413

Db 241 S 241

RESULT 11

Q9VLUS

ID Q9VLUS PRELIMINARY; PRT; 409 AA.

AC Q9VLUS;

DT 01-MAY-2000 (T-EMBLrel. 13, Created)

DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)

DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)

DE CG7221 protein (LD03827p).

GN CG7221.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkeley;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celnikier S.B., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Chertay J.M., Cusley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusekern D.R., Pacleb J.M.,  
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195 (2000).  
RN [2]  
RN SEQUENCE FROM N.A.  
RC STRAIN=Berkeley;  
RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,  
RA Champe W., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
RA George R., Gonzalez M., Guarin H., Grommler B., Li P., Liao G.,  
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,  
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
RA Celniker S.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
CC (SDR) FAMILY.  
DR EMBL; AE003619; AAF52587.1; -;  
DR EMBL; AY119574; AAM50228.1; -;  
DR HSP; Q13526; 1PIN.  
DR FlyBase; FBgn0031972; CG7221.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR InterPro; IPR002198; ADH short.  
DR InterPro; IPR001202; WW\_Rsp5\_WWP.  
DR Pfam; PF00106; adh\_short; 1.  
DR Pfam; PF00397; WW\_2.  
DR PROSITE; PS50020; WW\_DOMAIN\_2; 2.  
DR Oxidoreductase.  
SQ SEQUENCE 409 AA; 46510 MW; 7893BF4C39A7454C CRC64;

Query Match 46.4%; Score 1023; DB 5; Length 409;  
Best Local Similarity 49.5%; Pred. No. 2e-77;  
Matches 200; Conservative 68; Mismatches 126; Indels 10; Gaps 6;

QY 9 LDDTDSDELPGMEERTTKDGVVYANHTBEKTOWEHPKTKGRKRVAGDLPYGWEQETD 68  
Db 4 LPDTSDELPGMEERATDDGTVCVYVQNGKTSQWTHPTGRSKRIITGELPLGWEKYD 63

QY 69 ENGQVF-FVDHINKRTYLDRLAFTVDNDPTK--PTTRQRYDGSSTAMEILQGRDFTGK 125  
Db 64 BQGRFMPFNKETQORTNVDPRLAFVVEE-PTQNVAVQRQFDSCTALQVLHGKDLHGR 122

QY 126 VVVVTGANGSGGFPAKFAHGAHVILACRNMAPASAVSRILEEWHKAVEA--MTLD 183  
Db 123 TALITGANGCIGYETARSLAHHGCEIIFACNRNSAEAAIRIAQERPAARCRKFAALD 182

QY 184 LALLRSVQHFAEAFKAKNVPLHLVLCNAATFALPWSLTNDGLTTFQVNHHLGHFYLVL 243  
Db 183 LSSLSRVQRFVEEIKQSVSHIDYLLNLNAGVFPALPYTRTVDGLTTFQVSHLSHFYLTQL 242

QY 244 QDVLCRSAPARVIVSSSHRFTDINDSLGKLDLFSRLSPTRKNDYWMALAYNRSKLNILF 303  
Db 243 ETLF--DYKTRIIIVSSSEHFRANL--EVENLAVHLSPPPEKYWSMMAYNNAKLNVLF 298

QY 304 SNEHLRRLSPRGVTSNAVHPGNMYSNTHRSWVYTLFTLARPPTKSMQCGAATTVYCA 363  
Db 299 AQELAQRWKQGISVFSLHPGNVSSDLSRNFWFVRLFLFAIVRPPTKSLQQAASIIYCA 358

QY 364 AVEPELEGNGYFNCCRCMPSEPAQSEETARTLWALSERLQIE 407  
Db 359 TANELTGLSLGYFNCCFCFCEFSKLSKSAALQQLWKLSENLIAE 402

RESULT 12  
Q9NRK1 PRELIMINARY; PRT; 191 AA.  
AC Q9NRK1;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Fragile 16D oxido reductase (Fragment).  
GN FOR.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=20320695; PubMed=10861292;  
RA Ried K., Finnis M., Hobson L., Mangelsdorf M., Dayan S.,  
RA Nancarrow J.K., Woolatt E., Kremmidiotis G., Gardner A., Venter D.,  
RA Baker E., Richards R.I.;  
RT "Common chromosomal fragile site FRA16D sequence: Identification of the  
RT FOR gene spanning FRA16D and homozygous deletions and translocation  
RT breakpoints in cancer cells.";  
RL Hum. Mol. Genet. 9:1651-1663 (2000).  
DR EMBL; AF217430; AAF78197.1; -;  
DR EMBL; AF217491; AAF78197.1; JOINED.  
DR EMBL; AF217492; AAF78197.1; JOINED.  
DR GO; GO:0016491; F:oxidoreductase activity; NAS.  
FT NON TER 1  
SQ SEQUENCE 191 AA; 21721 MW; 030FC7E106D86B55 CRC64;

Query Match 42.8%; Score 942; DB 4; Length 191;  
Best Local Similarity 99.4%; Pred. No. 4.3e-71;  
Matches 179; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 173 HKAKVEAMTLDLALLRSVQHFAEAFKAKNVPLHLVLCNAATFALPWSLTNDGLTTFQVN 232  
Db 1 HKAKVETMTLDLALLRSVQHFAEAFKAKNVPLHLVLCNAATFALPWSLTNDGLTTFQVN 60

QY 233 HLGHFYLVLQDLVLCRSAPARVIVSSSEHFRFTDINDSLGKLDLFSRLSPTRKNDYWMALA 292  
Db 61 HLGHFYLVLQDLVLCRSAPARVIVSSSEHFRFTDINDSLGKLDLFSRLSPTRKNDYWMALA 120

QY 293 YNRSKLCNLFESNEHLRRLSPRGVTSNAVHPGNMYSNTHRSWVYTLFTLARPPTKSM 352  
Db 121 YNRSKLCNLFESNEHLRRLSPRGVTSNAVHPGNMYSNTHRSWVYTLFTLARPPTKSM 180

RESULT 13  
Q9NPC9 PRELIMINARY; PRT; 189 AA.  
AC Q9NPC9;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE FOR III protein (Truncated WW-domain oxidoreductase) (WW domain-  
DE containing oxidoreductase isoform FORIII).

```
GN WWOX.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20320695; PubMed=10861292;
RA Ried K., Finniss M., Hobson L., Mangelsdorf M., Dayan S., Venter D.,
RA Nancarrow J.K., Woolatt E., Kremmidiotis G., Gardner A.,
RA Baker E., Richards R.I.;
RT "Common chromosomal fragile site FRA16D sequence: Identification of the
RT FOR gene spanning FRA16D and homozygous deletions and translocation
RT breakpoints in cancer cells.";
RL Hum. Mol. Genet. 9:1651-1663(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX Chang N.-S., Pratt N., Carey G., Heath J.;
RA "A Novel WW Domain Oxidoreductase Enhances Tumor Necrosis Factor-
RT Mediated Cell Death and is a Partner of p53 in Apoptosis.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX PubMed=11572989;
RA Paige A.J.W., Taylor K.J., Taylor C., Hillier S.G., Farrington S.,
RA Scott D., Porteous D.J., Smyth J.F., Gabra H., Watson J.E.V.;
RT "WWOX: A candidate tumor suppressor gene involved in multiple tumor
RT types.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:11417-11422(2001).
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
DR EMBL; AF227528; AAF82055.1; -.
DR EMBL; AF187015; AAF31694.1; -.
DR EMBL; AF325429; AAL05450.1; -.
DR EMBL; AF325423; AAL05450.1; JOINED.
DR EMBL; AF325424; AAL05450.1; JOINED.
DR EMBL; AF325425; AAL05450.1; JOINED.
DR EMBL; AF325426; AAL05450.1; JOINED.
DR EMBL; AF325427; AAL05450.1; JOINED.
DR HSSP; Q13526; 1PIN.
DR GO; GO:0016491; F:oxidoreductase activity; NAS.
DR InterPro; IPR002198; ADH short.
DR InterPro; IPR001202; WW_Rep5_WWP.
DR Pfam; PF00106; adh short; 1.
DR Pfam; PF00397; WW; 2.
DR SMART; SM00456; WW; 2.
DR PROSITE; PS01159; WW DOMAIN 1; 2.
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
KW Oxidoreductase.
SQ SEQUENCE 189 AA; 21559 MW; 82155A9AD7C824C7 CRC64;

Query Match 42.1%; Score 927; DB 4; Length 189;
Best Local Similarity 100.0%; Pred. No. 7.7e-70;
Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAALRYAGLDLTDSEDELPFGWEERTTKDGVVYANHTTEKTQWEHPKTKRKRKRVAGDLP 60
Db 1 MAALRYAGLDLTDSEDELPFGWEERTTKDGVVYANHTTEKTQWEHPKTKRKRKRVAGDLP 60

Qy 61 YGWEQETDENGQVFFVDHINKRTTYLDPRLAFTVDDNPTKPTTRQRYDGSSTTAMEILQGR 120
Db 61 YGWEQETDENGQVFFVDHINKRTTYLDPRLAFTVDDNPTKPTTRQRYDGSSTTAMEILQGR 120

Qy 121 DFTGKVVVVTGANGSGIGFETAKSFALHGAHVILACRNARASEAVSRILEEW 172
Db 121 DFTGKVVVVTGANGSGIGFETAKSFALHGAHVILACRNARASEAVSRILEEW 172

RESULT 14
Q96KM3 ID Q96KM3 PRELIMINARY; PRT; 213 AA.
AC Q96KM3
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
```

```
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Aberrant WW domain-containing oxidoreductase.
GN WWOX.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11572989;
RA Paige A.J.W., Taylor K.J., Taylor C., Hillier S.G., Farrington S.,
RA Scott D., Porteous D.J., Smyth J.F., Gabra H., Watson J.E.V.;
RT "WWOX: A candidate tumor suppressor gene involved in multiple tumor
RT types.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:11417-11422(2001).
DR EMBL; AF325432; AAL05451.1; -.
DR EMBL; AF325423; AAL05451.1; JOINED.
DR EMBL; AF325424; AAL05451.1; JOINED.
DR EMBL; AF325425; AAL05451.1; JOINED.
DR EMBL; AF325426; AAL05451.1; JOINED.
DR EMBL; AF325433; AAL05451.1; JOINED.
DR InterPro; IPR001202; WW_Rep5_WWP.
DR Pfam; PF00397; WW; 2.
DR SMART; SM00456; WW; 2.
DR PROSITE; PS01159; WW DOMAIN 1; 2.
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
SQ SEQUENCE 213 AA; 23868 MW; A21054FF8214CC7C CRC64;

Query Match 34.1%; Score 750.5; DB 4; Length 213;
Best Local Similarity 79.7%; Pred. No. 6.2e-55;
Matches 145; Conservative 5; Mismatches 7; Indels 25; Gaps 2;

Qy 1 MAALRYAGLDLTDSEDELPFGWEERTTKDGVVYANHTTEKTQWEHPKTKRKRKRVAGDLP 60
Db 1 MAALRYAGLDLTDSEDELPFGWEERTTKDGVVYANHTTEKTQWEHPKTKRKRKRVAGDLP 60

Qy 61 YGWEQETDENGQVFFVDHINKRTTYLDPRLAFTVDDNPTKPTTRQRYDGSSTTAMEILQGR 120
Db 61 YGWEQETDENGQVFFVDHINKRTTYLDPRLAFTVDDNPTKPTTRQRYDGSSTTAMEILQGR 120

Qy 121 DFTGKVVVVTGANGSGIGFETAKSFALHGAHVILACRNARASEAVSRILEEWKAKYKAEAM 180
Db 121 DFTGKVVVVTGANGSGIG-----KASCHVGRTLKH---TRVEEL 155

Qy 181 TL 182
Db 156 SL 157

RESULT 15
Q96RF2 ID Q96RF2 PRELIMINARY; PRT; 311 AA.
AC Q96RF2
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE WWOXdelta5-8
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bednarek A.K., Keck-Waggoner C.L., Daniel R.L., Laflin K.J.,
RA Kluguchi K., Brenner A.J., Aldaz C.M.;
RT "WWOX, the FRA16D gene, behaves as a suppressor of tumor growth.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF395124; AAK81728.1; -.
DR InterPro; IPR001202; WW_Rep5_WWP.
DR Pfam; PF00397; WW; 2.
DR SMART; SM00456; WW; 2.
DR PROSITE; PS01159; WW_DOMAIN_1; 2.
```

DR PROSITE: PS50020: WW DOMAIN 2; 2;  
SQ SEQUENCE 311 AA; 35041 MW; 8EF6B9823F90C9F7 CRC64;  
Query Match 33.9%; Score 746.5; DB 4; Length 311;  
Best Local Similarity 90.3%; Pred. No. 2.3e-54;  
Matches 140; Conservative 3; Mismatches 5; Indels 7; Gaps 2;  
QY 1 MAALRYAGLDDTDESELPDGGWEERTTKDGVVYVYANHTTEKTQWEHPKTGKRKRVAGDLP 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1 MAALRYAGLDDTDESELPDGGWEERTTKDGVVYVYANHTTEKTQWEHPKTGKRKRVAGDLP 60  
QY 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDDNPTKPTTQRQYDGSSTAMEILQGR 120  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDDNPTKPTTQRQYDGSSTAMEILQGR 120  
QY 121 DFTGKVVVVTGANGSGIGFETAKSPALHGAHVILAC 155  
Db |||||||||||||||||||:|:|:|  
121 DFTGKVVVVTGANGSGIA-----TGSCH--HRVLCC 148

Search completed: May 24, 2004, 10:01:37  
Job time : 47 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: May 24, 2004, 21:49:29 ; Search time 8833 Seconds  
(without alignments)  
11109.321 Million cell updates/sec  
Title: US-09-978-318B-2  
Perfect score: 2264  
Sequence: 1 gcagcgccagcgtagcg.....aaaaaaaaaaaaaaaaaaaa 2264

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 5940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :	GenEmbl.*																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
------------	-----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

ALIGNMENTS

RESULT 1  
AF211943  
LOCUS AF211943 2264 bp mRNA linear PRI 05-MAY-2000  
DEFINITION Homo sapiens WW domain-containing protein WWOX mRNA, complete cds.  
ACCESSION AF211943  
VERSION AF211943.1 GI:6729682  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2264)  
AUTHORS Bednarek,A.K., Lafin,K.J., Daniel,R.L., Liao,Q., Hawkins,K.A. and Aldaz,C.M.  
TITLE WWOX, a novel WW domain-containing protein mapping to human

Pred. No. is the number of results predicted by chance to have a

chromosome 16q23.3-24.1, a region frequently affected in breast

Cancer Res. 60 (8), 2140-2145 (2000)

## JOURNAL

MEDLINE

20246348

PUBMED

10786676

REFERENCE

2 (bases 1 to 2264)

AUTHORS

Bednarek A.K. and Aldaz, C.M.

TITLE

Direct Submission

JOURNAL

Submitted (06-DEC-1999)

Anderson Cancer Center, Science Park-Research Division, Park Road

1C POB389, Smithville, TX 78957, USA

Location/Qualifiers

1. .2264

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/chromosome="16"

/map="16q23.3-q24.1"

1. .125

126. .1370

/note="putative hydroxysteroid dehydrogenase"

/codon\_start=1

/product="WW domain-containing protein WWOX"

/protein\_id="AAF27049.1"

/db\_xref="GI:6729683"

/translation="MAALRVAGLDDTDEDELPPWEERTTKDGMVYVYANHTREKTOW

EHPKTRKRVAGDLPVQWQSTDEQVFFVDHINKRTYLDPRLAFTYDDNPKPT

TRQYDSTAMEILQQRDFTGVVVVTGANSIGIFETAKSFALHGAHVILLACRNMAR

ASEAVRIIEEHKRYEAMTDLIALRSVQHFAEFAKKNVPVHLVIVCNAATFALPW

SLTKDGLFTFQVNLHGFYVQLQVLCRSAPARVIVVSSSEHRTFDINDSLGKLD

FSRLSPKNDYWMALAYNRSLCNILFSNELHRLSPRGVTSNAVHPGNMYSNIHRS

WVVYTLAFTLAPRPTKSMQQAATVYCAAPVPELGLGMYFNCCRCMPSPQAQSEE

TARTLWALSENLIQERLGSQSG"

1371. .2264

1371. .2264

2217. .2222

ORIGIN

Query Match

Best Local Similarity

Matches 2264; Conservative

100.08; Score 2264;

DB 9; Length 2264;

Indels 0; Gaps 0;

Mismatches 0; Pred. No. 0;

Query

1 GCAGTGGCGAGCGGTGAGCGGTGCGGGCCCGACGCGCGGGTCTCGTTTGGAGCGGAG 60

Db

1 GCAGTGGCGAGCGGTGAGCGGTGCGGGCCCGACGCGCGGGTCTCGTTTGGAGCGGAG 60

Query

61 TGAGTTCTGAGCAGTGGACCCGCGACGCGCGGCGATAGCGGGGCCAGGTGCTTCCACAGT 120

Db

61 TGAGTTCTGAGCAGTGGACCCGCGACGCGCGGCGATAGCGGGGCCAGGTGCTTCCACAGT 120

Query

121 CAGCCATGGCAGCGCTGCGCTACGCGGGGCTGGACGACGACGACGACGACGACGACGACG 180

Db

121 CAGCCATGGCAGCGCTGCGCTACGCGGGGCTGGACGACGACGACGACGACGACGACGACG 180

Query

181 CTCGGGCTGGGAGGAGAGAACCAACAGAGCGGCTGGGTTTACTACGCCAATCACACCG 240

Db

181 CTCGGGCTGGGAGGAGAGAACCAACAGAGCGGCTGGGTTTACTACGCCAATCACACCG 240

Query

241 AGGAGAGACTCAGTGGGAAACATCCAAAACCTGGAAGAAAGAAACGAGTGGCAGGAGATT 300

Db

241 AGGAGAGACTCAGTGGGAAACATCCAAAACCTGGAAGAAAGAAACGAGTGGCAGGAGATT 300

Query

301 TGCCATACGATGGGAAACAGAGAACTGATGAGACGAGCAAGTGTGTTGACCATTA 360

Db

301 TGCCATACGATGGGAAACAGAGAACTGATGAGACGAGCAAGTGTGTTGACCATTA 360

Query

361 TAAATAAAGAACCACTACTTTGGACCCCAAGACTGGCGTTTACTGTGGATGATATCCGA 420

Db

361 TAAATAAAGAACCACTACTTTGGACCCCAAGACTGGCGTTTACTGTGGATGATATCCGA 420

Query

421 CCAAGCCACCAACCCCGGAAAGATACGACGCGGACGACCACTGCGCATGGAATTCFCAGG 480

Db

421 CCAAGCCACCAACCCCGGAAAGATACGACGCGGACGACCACTGCGCATGGAATTCFCAGG 480

481 GCGGGGATTTCACTGGCAAAAGTGGTGTGTGTCTCACTGGAGCTAATTCAGGAATAGGGTTGG 540

481 GCGGGGATTTCACTGGCAAAAGTGGTGTGTGTCTCACTGGAGCTAATTCAGGAATAGGGTTGG 540

541 AAACCGCCAAAGTCTTTTGGCCCTCATGGTGCACATGTGATCTTGGCCCTGACGAAACATGG 600

541 AAACCGCCAAAGTCTTTTGGCCCTCATGGTGCACATGTGATCTTGGCCCTGACGAAACATGG 600

601 CAAGGCGAGTGAAGCAGTGTCCGCTCTGCTCGTAGCGTGCAGCAATTTTGTCTGAAGCATTAAG 660

601 CAAGGCGAGTGAAGCAGTGTCCGCTCTGCTCGTAGCGTGCAGCAATTTTGTCTGAAGCATTAAG 660

661 CAATGACCTTGGACCTCGCTCTGCTCGTAGCGTGCAGCAATTTTGTCTGAAGCATTAAG 720

661 CAATGACCTTGGACCTCGCTCTGCTCGTAGCGTGCAGCAATTTTGTCTGAAGCATTAAG 720

721 CCAAGAAATGTGCTTCTTCAATGTGCTTCAATGTGCTTCAATGTGCTTCAATGTGCTTCAATGTGCT 780

721 CCAAGAAATGTGCTTCTTCAATGTGCTTCAATGTGCTTCAATGTGCTTCAATGTGCTTCAATGTGCT 780

781 GTCTCACCAAAAGATGGCTCTGAGACCACTTTTCAAGTGAATCATCTGGGGCACTTCTTACC 840

781 GTCTCACCAAAAGATGGCTCTGAGACCACTTTTCAAGTGAATCATCTGGGGCACTTCTTACC 840

841 TTGTCCAGCTCCTCCAGATGTTTGTGCGCTCAGCTCCTGCGCGGTCTCAATGTGCTTCAATGTGCT 900

841 TTGTCCAGCTCCTCCAGATGTTTGTGCGCTCAGCTCCTGCGCGGTCTCAATGTGCTTCAATGTGCT 900

901 CCTCAGAGTCCCATCGATTTACAGATTTAAACGATCTCTTGGGAAACATGCACTTCACTGTC 960

901 CCTCAGAGTCCCATCGATTTACAGATTTAAACGATCTCTTGGGAAACATGCACTTCACTGTC 960

961 GCCTCTCTCCAAACAAAACGACATTTTGGGGGATGCTGGCTTAAACAGGTCCTCAAGCTCT 1020

961 GCCTCTCTCCAAACAAAACGACATTTTGGGGGATGCTGGCTTAAACAGGTCCTCAAGCTCT 1020

1021 GCAACATCTCTTCTCCAAAGAGCTGACCGTCTGCTCTCCCAACGCGGGGTCTCAAGCTCT 1080

1021 GCAACATCTCTTCTCCAAAGAGCTGACCGTCTGCTCTCCCAACGCGGGGTCTCAAGCTCT 1080

1081 ACGAGTGCATCTCGGAAATATGATGATCTTCAACATTCATCGCAGCTGCTGGGTGTACA 1140

1081 ACGAGTGCATCTCGGAAATATGATGATCTTCAACATTCATCGCAGCTGCTGGGTGTACA 1140

1141 CACTGCTGTTTACCTCTGCGAGGCGCTTTCACCAAGTCCATGCAACAGGAGCTGCGCACT 1200

1141 CACTGCTGTTTACCTCTGCGAGGCGCTTTCACCAAGTCCATGCAACAGGAGCTGCGCACT 1200

1201 CCGTGTACTGTGTCTGCTCCAGAACTGGAGGCTCTGGGAGGATGTACTTCAACAACT 1260

1201 CCGTGTACTGTGTCTGCTCCAGAACTGGAGGCTCTGGGAGGATGTACTTCAACAACT 1260

1261 GCTGCGGTGATGCTCTCCAGAACTGGAGGCTCTGGGAGGATGTACTTCAACAACT 1320

1261 GCTGCGGTGATGCTCTCCAGAACTGGAGGCTCTGGGAGGATGTACTTCAACAACT 1320

1321 CGCTCAGGAGGCTGATCCAGAACTGGAGGCTCTGGGAGGATGTACTTCAACAACT 1380

1321 CGCTCAGGAGGCTGATCCAGAACTGGAGGCTCTGGGAGGATGTACTTCAACAACT 1380

1381 GAGCGGATGGGCAACACACCGCTGT 1440

1381 GAGCGGATGGGCAACACACCGCTGT 1440

1441 GCCCTTCCAAATGTGCTTCCAAACAGAGTCCGAGAGTAAAGGAAATTAAGAGCAGTCA 1500

1441 GCCCTTCCAAATGTGCTTCCAAACAGAGTCCGAGAGTAAAGGAAATTAAGAGCAGTCA 1500

1501 CAAAGAGTGAAGAAATCTTAAGTACCAATGGAGAGGAGAAATTCCTGGGGTAAAGTATC 1560

1501 CAAAGAGTGAAGAAATCTTAAGTACCAATGGAGAGGAGAAATTCCTGGGGTAAAGTATC 1560

1561 ACTTTTCTGGGGCTGGGCTAGGCATAGTCTCTTTTGTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620





1 CAGTGCAGGCGGTGAGCGGTGCGGCCCGCAGCGCGCGGGTCTCTGTTTGGAGCGGGAGT 60  
62 GAGTTCTGAGCGAGTGGACCCCGGAGCGGGCGATAGGGGGGCGAGGTGCTTCCACAGTC 121  
61 GAGTTCTGAGCGAGTGGACCCCGGAGCGGGCGATAGGGGGGCGAGGTGCTTCCACAGTC 120  
122 AGCCATGTGAGCGCTCGCTTACCGCGGGCTGGACGACACGAGCAGTGGAGGACGAGTGGC 181  
121 AGCCATGTGAGCGCTCGCTTACCGCGGGCTGGACGACACGAGCAGTGGAGGACGAGTGGC 180  
182 TCCGGGCTGGAGGAGAGAAACCAAGGACCGCTGGGTTTACTAGCCCAATCACACCGA 241  
181 TCCGGGCTGGAGGAGAGAAACCAAGGACCGCTGGGTTTACTAGCCCAATCACACCGA 240  
242 GGAAGAATCAGTGGGAACATCCAAAACCTGAAAACGAAAAGAAAACGAGTGGCAGAGATTT 301  
241 GGAAGAATCAGTGGGAACATCCAAAACCTGAAAACGAAAAGAAAACGAGTGGCAGAGATTT 300  
302 GCCATACGGATGGAAACAGAAATCATGAGAACGGACAGTGTGTTTGTTCACCATAT 361  
301 GCCATACGGATGGAAACAGAAATCATGAGAACGGACAGTGTGTTTGTTCACCATAT 360  
362 AAATAAAGAACCACTTACTTGGACCCAGACTGGCGTTTACTGTGGATGATTAATCCGAC 421  
361 AAATAAAGAACCACTTACTTGGACCCAGACTGGCGTTTACTGTGGATGATTAATCCGAC 420  
422 CAAGCCAAACCCCGCAAGATACGACGCGACACCACTGCCATGGAATTTCTCCAGG 481  
421 CAAGCCAAACCCCGCAAGATACGACGCGACACCACTGCCATGGAATTTCTCCAGG 480  
482 CCGGGAATTCAGTGGCAAGTGTGTGGTCACTGGAGCTAAATCAGGAATAGGGTTGGA 541  
481 CCGGGAATTCAGTGGCAAGTGTGTGGTCACTGGAGCTAAATCAGGAATAGGGTTGGA 540  
542 AACCGCAAGTCTTTTGGCCCTCATGGTGCACATGTGATCTGGCCCTGCAGGAACATGGC 601  
541 AACCGCAAGTCTTTTGGCCCTCATGGTGCACATGTGATCTGGCCCTGCAGGAACATGGC 600  
602 AAGGGGAGTGAAGCAGTGTACGCAATTTAGAAAGATGGCATAAAGCCGAGGTAGAGC 661  
601 AAGGGGAGTGAAGCAGTGTACGCAATTTAGAAAGATGGCATAAAGCCGAGGTAGAGC 660  
662 AATGACCTGGACCTCGCTCTGCTCGTAGCGGTCAGCATTTTGTCTGAAGCATTTCAAGGC 721  
661 AATGACCTGGACCTCGCTCTGCTCGTAGCGGTCAGCATTTTGTCTGAAGCATTTCAAGGC 720  
722 CAAGAATGTGCCCTTTCAATGTCTTGTGTGCACACGAGCAACTTTTGTCTACCCCTGGAG 781  
721 CAAGAATGTGCCCTTTCAATGTCTTGTGTGCACACGAGCAACTTTTGTCTACCCCTGGAG 780  
782 TCTCACAAGATGGCGCTGGAGACCACTTTCAAGTGAATCATCTGGGGCACTTCTACCT 841  
781 TCTCACAAGATGGCGCTGGAGACCACTTTCAAGTGAATCATCTGGGGCACTTCTACCT 840  
842 TGTCCAGCTCTCCAGAGTGTGTTTGTGCGCTCAGCTCCCTGCGGTGTCAATTTGGTGTCTC 901  
841 TGTCCAGCTCTCCAGAGTGTGTTTGTGCGCTCAGCTCCCTGCGGTGTCAATTTGGTGTCTC 900  
902 CTCAGAGTCCCATCGATTTACAGATATTAACGACTCCCTGGGAAAACCTGACATTCAGTGC 961  
901 CTCAGAGTCCCATCGATTTACAGATATTAACGACTCCCTGGGAAAACCTGACATTCAGTGC 960  
962 CCTCTCACAAGAACCAACTATTGGGCGATGCTGGCTTATAACAGGTCCAGCTCTG 1021  
961 CCTCTCTCACAAGAACCAACTATTGGGCGATGCTGGCTTATAACAGGTCCAGCTCTG 1020  
1022 CAACATCTCTCTCCAAACAGATGACCGTGCCTCTCCCAACGCGGGGTCAAGTCCGAA 1081  
1021 CAACATCTCTCTCCAAACAGATGACCGTGCCTCTCCCAACGCGGGGTCAAGTCCGAA 1080  
1082 CGCAGTGCATCTCGAAATATGATGACTCCAACTCATCGCAGCTGTGGGTGTACAC 1141  
1081 CGCAGTGCATCTCGAAATATGATGACTCCAACTCATCGCAGCTGTGGGTGTACAC 1140

1142 ACTGCTGTTTACCTTGGCGAGGCGCTTTTCAACAAGTCCATGCAACAGAGGAGCTGCACCAC 1201  
1141 ACTGCTGTTTACCTTGGCGAGGCGCTTTTCAACAAGTCCATGCAACAGAGGAGCTGCACCAC 1200  
1202 CGTGTACTGTGCTGTGCTCCCGAACTGGAGGCTCTGGGAGGATGTACTTCAACAACATG 1261  
1201 CGTGTACTGTGCTGTGCTCCCGAACTGGAGGCTCTGGGAGGATGTACTTCAACAACATG 1260  
1262 CTGCGCTGCATGCCCTCACAGAACTCAGACGGAAGACGCGCCCGGACCCCTGTGGGC 1321  
1261 CTGCGCTGCATGCCCTCACAGAACTCAGACGGAAGACGCGCCCGGACCCCTGTGGGC 1320  
1322 GCTCAGCGAGAGCTCATCCAGAACTGACGCGCTTGCACGCTCCGCTTAAGTGAGGCTCAG 1381  
1321 GCTCAGCGAGAGCTCATCCAGAACTGACGCGCTTGCACGCTCCGCTTAAGTGAGGCTCAG 1380  
1382 AGCGATGGGCAACACACCGCCCTGTGTGTGCTCCCTCACGCAAGTGCAGGGGTGGG 1441  
1381 AGCGATGGGCAACACACCGCCCTGTGTGTGCTCCCTCACGCAAGTGCAGGGGTGGG 1440  
1442 CCCCTTCCAAATGTCCTCCACACAGATCCGCAAGTAAAGGAAATAGACGAGTCAC 1501  
1441 CCCCTTCCAAATGTCCTCCACACAGATCCGCAAGTAAAGGAAATAGACGAGTCAC 1500  
1502 AACAGAGTGAATAATCTTAAGTACCAATGGAAGCAGGGAATTCCTGGGCTAAAGTATCA 1561  
1501 AACAGAGTGAATAATCTTAAGTACCAATGGAAGCAGGGAATTCCTGGGCTAAAGTATCA 1560  
1562 CTTTCTCGGGCTGGGCTAGGCATAGTCTCTTTGCTTTCTGCTGGTGGCTGTGTTGAAA 1621  
1561 CTTTCTCGGGCTGGGCTAGGCATAGTCTCTTTGCTTTCTGCTGGTGGCTGTGTTGAAA 1620  
1622 GTAAAACTGCTTGTGTGTGTAGGTTCCGTATCTCCTGGAAGAACACAGCAATTTCTCT 1681  
1621 GTAAAACTGCTTGTGTGTGTAGGTTCCGTATCTCCTGGAAGAACACAGCAATTTCTCT 1680  
1682 TTTCTTTTACTGTTATAGAAATAGCTGAGTCCCTCGTCCCATCCAGTACCAACGCGC 1741  
1681 TTTCTTTTACTGTTATAGAAATAGCTGAGTCCCTCGTCCCATCCAGTACCAACGCGC 1740  
1742 CACACTGCAGCGGGGCTGGCTTTCTCCTACTTAGGGAAGAAAAGCAAGTGTTCAC 1801  
1741 CACCACTGCAGCGGGGCTGGCTTTCTCCTACTTAGGGAAGAAAAGCAAGTGTTCAC 1800  
1802 GCTCCTTGTGCTGATTCAGAGAGATAATGTTTCAATCTCCTGACCAAGCTGAGCC 1861  
1801 GCTCCTTGTGCTGATTCAGAGAGATAATGTTTCAATCTCCTGACCAAGCTGAGCC 1860  
1862 AGCTTAGCAACTGCTGGGAGACAAATCTCAGAACTTGTCCCGAGCCAGTGAGGATGACA 1921  
1861 AGCTTAGCAACTGCTGGGAGACAAATCTCAGAACTTGTCCCGAGCCAGTGAGGATGACA 1920  
1922 GTGACACCCAGAGGGAGTAGAATACGAGAACTACAGGTGGCAAGTACTTGTCAATAGA 1981  
1921 GTGACACCCAGAGGGAGTAGAATACGAGAACTACAGGTGGCAAGTACTTGTCAATAGA 1980  
1982 CTCCTTTGCTTAATGCTATGCAAAAAATTTCTTAGAGATATTAACAAATTTTCAATCAT 2041  
1981 CTCCTTTGCTTAATGCTATGCAAAAAATTTCTTAGAGATATTAACAAATTTTCAATCAT 2040  
2042 TCCCTTAGATACCTTTGAAAGCAGGAAGGAGCTATATATCTTAAGAAATACAGAGATAT 2101  
2041 TCCCTTAGATACCTTTGAAAGCAGGAAGGAGCTATATATCTTAAGAAATACAGAGATAT 2100  
2102 TTTGGGGGCGAGAGATAAAACGTTAGTTAAATCCCTTTTGTCTGCTCAATCAAGTCTCAGT 2161  
2101 TTTGGGGGCGAGAGATAAAACGTTAGTTAAATCCCTTTTGTCTGCTCAATCAAGTCTCAGT 2160  
2162 TCTCTTGTCTTTCAATGTATCTTAACCTCTGCTGCTCGCTCGCATCTTATGCTTAAATA 2221  
2161 TCTCTTGTCTTTCAATGTATCTTAACCTCTGCTGCTCGCTCGCATCTTATGCTTAAATA 2220

QY 2222 AAGAACATGCTTGAATATCAAAAAA 2256  
Db 2221 AAGAACATGCTTGAATATCAAAAAA 2255

RESULT 4  
BD156478  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BD156478 2256 bp DNA linear PAT 17-JAN-2003  
Primer for synthesizing full-length cDNA and use thereof.  
BD156478  
BD156478.1 GI:27862236  
JP 2002191363-A/11321.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1. (bases 1 to 2256)  
Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,  
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.  
Primer for synthesizing full-length cDNA and use thereof  
Patent: JP 2002191363-A 11321 09-JUL-2002;  
HELIX RESEARCH INSTITUTE  
OS Homo sapiens (human)  
PN JP 2002191363-A/11321  
PD 09-JUL-2002  
PF 28-JUL-2000 JP 2000280990  
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU  
PI SAITO,  
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,  
PI KEIICHI NAGAI, TETSUO OTSUKI  
PC  
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC  
10,  
C12N5/09, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/09 CC  
PC, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC  
Primer for synthesizing full-length cDNA and use thereof FH Key  
Location/Qualifiers  
FT CDS (125). (1366).  
source  
1. 2256  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 99.5%; Score 2253.4; DB 6; Length 2256;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CAGTGCAGCGCTGAGCGCTCGGCGCCCGACGCGCGCGGTCTCGTTTGGAGCGGAGT 61  
Db 1 CAGTGCAGCGCTGAGCGCTCGGCGCCCGACGCGCGCGGTCTCGTTTGGAGCGGAGT 60

QY 62 GAGTTCCTGAGCGAGTGGACCGCGGCGAGCGGCGGTAGGGGGCGAGTCCCTCCACAGTC 121  
Db 61 GAGTTCCTGAGCGAGTGGACCGCGGCGAGCGGCGGTAGGGGGCGAGTCCCTCCACAGTC 120

QY 122 AGCCATGGCAGCGCTGCGCTAGCGGGGCTGGACACACGACGAGTGGAGCGAGTGGCC 181  
Db 121 AGCCATGGCAGCGCTGCGCTAGCGGGGCTGGACACACGACGAGTGGAGCGAGTGGCC 180

QY 182 TCCGGCTGGAGGAGAGAACCAACGAGCGGCTGGTTTACTAGCCCAATCAACACCGA 241  
Db 181 TCCGGCTGGAGGAGAGAACCAACGAGCGGCTGGTTTACTAGCCCAATCAACACCGA 240

QY 242 GGAGAGACTCAGTGGGACATCCAAACCTGGAAACGAGTGGAGGAGATTT 301  
Db 241 GGAGAGACTCAGTGGGACATCCAAACCTGGAAACGAGTGGAGGAGATTT 300

QY 302 GCCATACGATGGGAACAAGAACTGATGAGAAACGACAAAGTGTGTTTGTGACCATAT 361  
Db 301 GCCATACGATGGGAACAAGAACTGATGAGAAACGACAAAGTGTGTTTGTGACCATAT 360

QY 362 AAATAAAGAACCACTACTTGGACCAAGACTGGCGTTTACTGTGGATGATAATCCGAC 421

Db 361 AAATAAAGAACCACTACTTGGACCAAGACTGGCGTTTACTGTGGATGATAATCCGAC 420  
QY 422 CAAGCCAAACCCCGCAAGATACGACGAGCACCCTGCGATGGAATTTCTCCAGGG 481  
Db 421 CAAGCCAAACCCCGCAAGATACGACGAGCACCCTGCGATGGAATTTCTCCAGGG 480

QY 482 CCGGATTTCACTGGCAAAAGTGTGTGCTACTGGAGCTAATTCAGGAATAGGGTTGA 541  
Db 481 CCGGATTTCACTGGCAAAAGTGTGTGCTACTGGAGCTAATTCAGGAATAGGGTTGA 540

QY 542 AACCCGCCAAGTCTTTTGGCCCTCATGGTGCAATGTGATCTTTGGCTGCAGGAACATGGC 601  
Db 541 AACCCGCCAAGTCTTTTGGCCCTCATGGTGCAATGTGATCTTTGGCTGCAGGAACATGGC 600

QY 602 AAGGCGGAGTGAAGCAGTGTGACGATTTTGAAGAATGGCATAAGCCCAAGTAGAAGC 661  
Db 601 AAGGCGGAGTGAAGCAGTGTGACGATTTTGAAGAATGGCATAAGCCCAAGTAGAAGC 660

QY 662 AATGACCCCTGGACCTCGCTCTCGCTAGCTGAGCAGATTTTGTGGAAGCATTCAAGGC 721  
Db 661 AATGACCCCTGGACCTCGCTCTCGCTAGCTGAGCAGATTTTGTGGAAGCATTCAAGGC 720

QY 722 CAAGAATGTGCTCTTTCATGTGCTTTGTGCAACGACCAACTTTTGTCTTACCTCGGAG 781  
Db 721 CAAGAATGTGCTCTTTCATGTGCTTTGTGCAACGACCAACTTTTGTCTTACCTCGGAG 780

QY 782 TCTCACCACCAAGATGGCTGGAGACACACCTTTCAAGTGAATCATCTGGGCACTTCTACCT 841  
Db 781 TCTCACCACCAAGATGGCTGGAGACACACCTTTCAAGTGAATCATCTGGGCACTTCTACCT 840

QY 842 TGTCCAGCTCTCTCCAGGATGTTTGTGCGCTCAGCTCTCTGCGCTGTCTATTGTGGTCTC 901  
Db 841 TGTCCAGCTCTCTCCAGGATGTTTGTGCGCTCAGCTCTCTGCGCTGTCTATTGTGGTCTC 900

QY 902 CTCAGAGTCCCATCGATTTACAGATATTAACGATCTCTTTGGGAAACTGGACTTCAGTGC 961  
Db 901 CTCAGAGTCCCATCGATTTACAGATATTAACGATCTCTTTGGGAAACTGGACTTCAGTGC 960

QY 962 CCTCTCTCCACAAAAACGACTATTTGGCGATGCTGGCTTATTAACAGTCCCAAGTCTG 1021  
Db 961 CCTCTCTCCACAAAAACGACTATTTGGCGATGCTGGCTTATTAACAGTCCCAAGTCTG 1020

QY 1022 CAACATCTCTCTCCACAGAGTGCACCGTCCCTCTCCCGACGCGGCTCACGTCGAA 1081  
Db 1021 CAACATCTCTCTCCACAGAGTGCACCGTCCCTCTCCCGACGCGGCTCACGTCGAA 1080

QY 1082 CGCAGTGCATCTCGAAATATGATGTATCTCCAAATTCATTCGACGTGGTGGGTGTACAC 1141  
Db 1081 CGCAGTGCATCTCGAAATATGATGTATCTCCAAATTCATTCGACGTGGTGGGTGTACAC 1140

QY 1142 ACTGCTGTGTTTACCTTGGCGAGGCTTTCACCAAGTCCATGCAACAGGAGTGCACCAAC 1201  
Db 1141 ACTGCTGTGTTTACCTTGGCGAGGCTTTCACCAAGTCCATGCAACAGGAGTGCACCAAC 1200

QY 1202 CGTGTACTGTGCTGTCTCCAGAACTGGAGGGTCTGGAGGGATGTACTTCAACAACTG 1261  
Db 1201 CGTGTACTGTGCTGTCTCCAGAACTGGAGGGTCTGGAGGGATGTACTTCAACAACTG 1260

QY 1262 CTGCGCTGCATGCGCTTCAACAGAGTCTCAGACGAAAGAGACGCGCCCGGACCCCTGTGGGC 1321  
Db 1261 CTGCGCTGCATGCGCTTCAACAGAGTCTCAGACGAAAGAGACGCGCCCGGACCCCTGTGGGC 1320

QY 1322 GCTCAGCGCAGAGGCTGTATCCCAAGACGCTTGGCAGCCAGTCCCGCTAAGTGGAGCTCAG 1381  
Db 1321 GCTCAGCGCAGAGGCTGTATCCCAAGACGCTTGGCAGCCAGTCCCGCTAAGTGGAGCTCAG 1380

QY 1382 AGCGGATGGGACACACACCGCGCTGTGTGTCTCCCTCAGCAAGTGCAGGCTGGG 1441  
Db 1381 AGCGGATGGGACACACACCGCGCTGTGTGTCTCCCTCAGCAAGTGCAGGCTGGG 1440

QY 1442 CCCCTTCCAAATGCTCCCTCCAAACAGATCCGCAAGATTAAGGAATTAAGAGCTGCTAC 1501

Db 1441 CCCCTTCCAAATGTCCTCCCAACACAGATCCGCAAGAGTAAGAAATAGAGCATCTAC 1500

QY 1502 AACAGAGTGAATAATCTTAAGTACCAATGGAAGCAGGGAATTCCTGGGGTAAAGTATCA 1561

Db 1501 AACAGAGTGAATAATCTTAAGTACCAATGGAAGCAGGGAATTCCTGGGGTAAAGTATCA 1560

QY 1562 CTTTCTCGGGCTGGGCTAGGCATAGGTCTCTTTGCTTTCTGTGTGTGGCTGTTGAAA 1621

Db 1561 CTTTCTCGGGCTGGGCTAGGCATAGGTCTCTTTGCTTTCTGTGTGTGGCTGTTGAAA 1620

QY 1622 GTAAAAACCTGCTGTGTGTAGGTTCGGTATCTCCTGTGAGAGCAGCAGCAATCTCT 1681

Db 1621 GTAAAAACCTGCTGTGTGTAGGTTCGGTATCTCCTGTGAGAGCAGCAGCAATCTCT 1680

QY 1682 TTCTTTTACTGTTATAGATAGCTGAGTCCCTCTGTCCTCCATCCAGTACCAACACGGC 1741

Db 1681 TTCTTTTACTGTTATAGATAGCTGAGTCCCTCTGTCCTCCATCCAGTACCAACACGGC 1740

QY 1742 CACCACTGACGCGGGGCTGGCTTCTCTACTATAGGGAAGAAAAAGCAAGTGTTCAT 1801

Db 1741 CACCACTGACGCGGGGCTGGCTTCTCTACTATAGGGAAGAAAAAGCAAGTGTTCAT 1800

QY 1802 GCTCCTTGTGCTATGATCCAGAGATATTTGTTTCAATCTGACCAAGTGTTCAGCC 1861

Db 1801 GCTCCTTGTGCTATGATCCAGAGATATTTGTTTCAATCTGACCAAGTGTTCAGCC 1860

QY 1862 AGCTTAGCAACTCTGGGAGAGCAAAATCTCAGAACCTTGTCCAGCCAGTGAGGATGACA 1921

Db 1861 AGCTTAGCAACTCTGGGAGAGCAAAATCTCAGAACCTTGTCCAGCCAGTGAGGATGACA 1920

QY 1922 GTGACACCCAGAGGAGTGAATACGAGAACTACAGGTGGCAAAAGTACTTGTCTATAGA 1981

Db 1921 GTGACACCCAGAGGAGTGAATACGAGAACTACAGGTGGCAAAAGTACTTGTCTATAGA 1980

QY 1982 CTCCTTTGCTATGCTATGCAAAAATCTCTTAGAGATTTATACAAATTTTCAATCAT 2041

Db 1981 CTCCTTTGCTATGCTATGCAAAAATCTCTTAGAGATTTATACAAATTTTCAATCAT 2040

QY 2042 TCCTTAGATACCTTGAAAGCAGGAGGAGCGGTATATATCTTAAGAAATACACAGGATAT 2101

Db 2041 TCCTTAGATACCTTGAAAGCAGGAGGAGCGGTATATATCTTAAGAAATACACAGGATAT 2100

QY 2102 TTTGGGGGCGAGAGATAAAGCTTATAGTTAATCCCTTTGCTGTCTCAATCACAGTCTCAGT 2161

Db 2101 TTTGGGGGCGAGAGATAAAGCTTATAGTTAATCCCTTTGCTGTCTCAATCACAGTCTCAGT 2160

QY 2162 TCTCTTGTCTTCACTGATTTAAACCTCTGCTGTGCTCGCTCCATCTGTTTATATA 2221

Db 2161 TCTCTTGTCTTCACTGATTTAAACCTCTGCTGTGCTCGCTCCATCTGTTTATATA 2220

QY 2222 AAGAACATGCTTGAATATCAAAAAAAGAAAAA 2256

Db 2221 AAGAACATGCTTGAATATCAAAAAAAGAAAAA 2255

RESULT 5  
AK027626  
LOCUS  
DEFINITION Homo sapiens cDNA FLJ14720 fis, clone NT2RP3001495, highly similar to Human oxidoreductase (HHCW56) mRNA.  
ACCESSION AK027626  
VERSION AK027626.1 GI:14042433  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H., Wagaesuma, M., Hosori, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takanashi, M., Chiba, Y., Ishida, S., Murakawa, K., Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K.,

# TITLE JOURNAL AUTHORS REFERENCE JOURNAL

## COMMENT

## FEATURES source

1. .2256  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="NT2RP3001495"  
/cell\_line="NT2"  
/clone\_lib="NT2RP3"  
/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction."

## ORIGIN

Query Match 99.5%; Score 2253.4; DB 9; Length 2256;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CAGTGGCGAGCGGTGAGCGGTGCGGCCCGGACCGCGCGCGGTCTCGTTGGAGCGGAGT 61

Db 1 CAGTGGCGAGCGGTGAGCGGTGCGGCCCGGACCGCGCGCGGTCTCGTTGGAGCGGAGT 60

QY 62 GAGTTCTCAGCGAGTGGACCGCGGACCGGCGGCGATAGGGGGCCAGTGCCTCCAGTCT 121

Db 61 GAGTTCTCAGCGAGTGGACCGCGGACCGGCGGCGATAGGGGGCCAGTGCCTCCAGTCT 120

QY 122 AGCCATGGCAGCGCTCGGTACGCGGGGCTGGACGACGACGACGAGTGGACGAGTGC 181

Db 121 AGCCATGGCAGCGCTCGGTACGCGGGGCTGGACGACGACGAGTGGACGAGTGC 180

QY 182 TCCGGGCTGGAGGAGAGAACCAACCAAGGACCGCTGGTTTACTAGCCCAATCACACCGA 241

Db 181 TCCGGGCTGGAGGAGAGAACCAACCAAGGACCGCTGGTTTACTAGCCCAATCACACCGA 240

QY 242 GGNAGAGACTCAGTGGGACATCCAAATCTGAAAGAGAAAGAGAGTGGCAGGAGATT 301

Db 241 GGNAGAGACTCAGTGGGACATCCAAATCTGAAAGAGAAAGAGAGTGGCAGGAGATT 300

QY 302 GCCATACCGATGGAAACAAAGAACTGATGAGAACGACGACGAGTGTCTTTCTGACCATAT 361

Db 301 GCCATACCGATGGAAACAAAGAACTGATGAGAACGACGACGAGTGTCTTTGACCATAT 360

QY 362 AATAAAGAACCACTTCTTGGACCCAGAGTGGCGTTTACTGTGGATGATATCCGAC 421

Db 361 AATAAAGAACCACTTCTTGGACCCAGAGTGGCGTTTACTGTGGATGATATCCGAC 420

QY 422 CAGCCCAACCCCGGCAAGATACGACGCGCAGCACCACTGCCATCGAATTTCCAGG 481

Db 421 CAGCCCAACCCCGGCAAGATACGACGCGCAGCACCACTGCCATCGAATTTCCAGG 480

QY 482 CCGGGATTTCTCTGGCAAAAGTGGTGTGTCATCTGGAGCTAAATTCAGGAAATAGGTTGA 541

Db 481 CCGGGATTTCTCTGGCAAAAGTGGTGTGTCATCTGGAGCTAAATTCAGGAAATAGGTTGA 540

QY 542 AACCGGCAAGTCTTTTGCCTCCATGTCATGTCATGTCATGTCATGTCATGTCATGTC 601

Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.  
NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 2256)  
Isogai, T. and Otsuki, T.  
Direct Submission  
Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

## Location/Qualifiers



Db 541 AACGCCAAGCTTTTGGCCCTCATGGTGACATGTGATCTTGGCCCTGAGCAACATGGC 600  
Qy 602 AAGGCGAGTGAAGCAGTGTACGCAATTTTGAAGATGGCATAAAGCCAAAGGTAGAGC 661  
Db 601 AAGGCGAGTGAAGCAGTGTACGCAATTTTGAAGATGGCATAAAGCCAAAGGTAGAGC 660  
Qy 662 AATGACCTTGAACCTCGCTCTGCTCCGTAGCGTGCAGCATTTTGTGAAGCATTTCAAGGC 721  
Db 661 AATGACCTTGAACCTCGCTCTGCTCCGTAGCGTGCAGCATTTTGTGAAGCATTTCAAGGC 720  
Qy 722 CAAGATGTGCTCTTCATGTGCTGTGTGCAACGAGCACTTTTGTCTACCTCGAG 781  
Db 721 CAAGATGTGCTCTTCATGTGCTGTGTGCAACGAGCACTTTTGTCTACCTCGAG 780  
Qy 782 TCTCACCAAGATGCGCTGGAGACCACTTTTCAAGTGAATCATCTGGGCGCACTTCACT 841  
Db 781 TCTCACCAAGATGCGCTGGAGACCACTTTTCAAGTGAATCATCTGGGCGCACTTCACT 840  
Qy 842 TGTCCAGCTCTTCCAGGATGTTTGTGCGGCTCAGCTCCTGCGGTGTCATTGTGGTCTC 901  
Db 841 TGTCCAGCTCTTCCAGGATGTTTGTGCGGCTCAGCTCCTGCGGTGTCATTGTGGTCTC 900  
Qy 902 CTCAGAGTCCCATGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 961  
Db 901 CTCAGAGTCCCATGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 960  
Qy 962 CCTCTCTCCAAACAAACAGCTATTGGGCGATGCTGGCTTATTAACAGGTCCAGCTCTG 1021  
Db 961 CCTCTCTCCAAACAAACAGCTATTGGGCGATGCTGGCTTATTAACAGGTCCAGCTCTG 1020  
Qy 1022 CAACATCTCTTCTCCAAACAGCTGCGCTGCGCTCTCCCAACGCGGGGTACGTCGAA 1081  
Db 1021 CAACATCTCTTCTCCAAACAGCTGCGCTGCGCTCTCCCAACGCGGGGTACGTCGAA 1080  
Qy 1082 CGCAGTGCATCTGGAATATGATGATCTCCAACTTATTCAGCAGCTGGTGGTGTAC 1141  
Db 1081 CGCAGTGCATCTGGAATATGATGATCTCCAACTTATTCAGCAGCTGGTGGTGTAC 1140  
Qy 1142 ACTGCTGTTTACCTTGGCGAGGCTTTCCAAAGTCCATGCAACAGGAGCTGCCACCAC 1201  
Db 1141 ACTGCTGTTTACCTTGGCGAGGCTTTCCAAAGTCCATGCAACAGGAGCTGCCACCAC 1200  
Qy 1202 CGTGTACTGTGCTGTCTCCCAAGACTGGAGGCTGGAGGAGTGTACTTCAACAACTG 1261  
Db 1201 CGTGTACTGTGCTGTCTCCCAAGACTGGAGGCTGGAGGAGTGTACTTCAACAACTG 1260  
Qy 1262 CTGCGCTGCTATGCTCCTCAGAGAGCTCAGAGCGAGGAGCGGCCCGGACCTGTGGC 1321  
Db 1261 CTGCGCTGCTATGCTCCTCAGAGAGCTCAGAGCGAGGAGCGGCCCGGACCTGTGGC 1320  
Qy 1322 GTCAGCGAGAGGCTGTATCCCAAGACGCTTGGCAGCGAGTCCGGCTAAGTGGAGCTCAG 1381  
Db 1321 GTCAGCGAGAGGCTGTATCCCAAGACGCTTGGCAGCGAGTCCGGCTAAGTGGAGCTCAG 1380  
Qy 1382 AGCGATGGGACACACACCGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1441  
Db 1381 AGCGATGGGACACACACCGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440  
Qy 1442 CCCCTTCCAAATGCTCCTCCAAACAGATCCGCAAGATTAAGGAATAAGAGCAGTCCAC 1501  
Db 1441 CCCCTTCCAAATGCTCCTCCAAACAGATCCGCAAGATTAAGGAATAAGAGCAGTCCAC 1500  
Qy 1502 AACAGAGTGAAATCTTAAAGTACCAATGGGAAGCAGGGAATCTCTGGGGTAAAGTATCA 1561  
Db 1501 AACAGAGTGAAATCTTAAAGTACCAATGGGAAGCAGGGAATCTCTGGGGTAAAGTATCA 1560  
Qy 1562 CTTTCTGGGCTGGGCTAGGATAGGTCTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 1621  
Db 1561 CTTTCTGGGCTGGGCTAGGATAGGTCTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 1620  
Qy 1622 GTAAAACTGCTGT 1681  
Db 1621 GTAAAACTGCTGT 1680

Qy 1682 TTCTTTTACTGTATAGATAGCTAGGTCCTCGTCCCATCCAGTCCAGTCCACACGGC 1741  
Db 1681 TTCTTTTACTGTATAGATAGCTAGGTCCTCGTCCCATCCAGTCCACACGGC 1740  
Qy 1742 CACCACTGACCGGGGCTGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1801  
Db 1741 CACCACTGACCGGGGCTGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1800  
Qy 1802 GCTCCTTGTCTGATTCAGAGGAGTAAATTTTCTTCTCTCTCTCTCTCTCTCTCTCTCT 1861  
Db 1801 GCTCCTTGTCTGATTCAGAGGAGTAAATTTTCTTCTCTCTCTCTCTCTCTCTCTCTCT 1860  
Qy 1862 AGCTTAGCAACTGCTGGGAGACAAATCTCAGAACTTGTCTCCAGCCAGTGGAGTAC 1921  
Db 1861 AGCTTAGCAACTGCTGGGAGACAAATCTCAGAACTTGTCTCCAGCCAGTGGAGTAC 1920  
Qy 1922 GTGACACCCAGGAGGATAGATACGAGAACTACAGGTTGGCAAGTACTTGTCTATAGA 1981  
Db 1921 GTGACACCCAGGAGGATAGATACGAGAACTACAGGTTGGCAAGTACTTGTCTATAGA 1980  
Qy 1982 CTCCTTTGTCTGATTCAGAGGAGTAAATTTTCTTCTCTCTCTCTCTCTCTCTCTCTCT 2041  
Db 1981 CTCCTTTGTCTGATTCAGAGGAGTAAATTTTCTTCTCTCTCTCTCTCTCTCTCTCTCT 2040  
Qy 2042 TCCTTAGATACCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2101  
Db 2041 TCCTTAGATACCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2100  
Qy 2102 TTTGGGGGCGCAGAGATTAATAAAGCTTGTAGTAAATCCCTTGTCTGTCAATTCACAGTCTCAGT 2161  
Db 2101 TTTGGGGGCGCAGAGATTAATAAAGCTTGTAGTAAATCCCTTGTCTGTCAATTCACAGTCTCAGT 2160  
Qy 2162 TCTCTTGTCTTTCATGATGCTTAAACCTCTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 2221  
Db 2161 TCTCTTGTCTTTCATGATGCTTAAACCTCTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 2220  
Qy 2222 AAGAACATGCTTGAATATCAAAAAAATAAAAAA 2256  
Db 2221 AAGAACATGCTTGAATATCAAAAAAATAAAAAA 2255

## RESULT 6

AF227527 2219 bp mRNA linear PRI 06-JUL-2000  
LOCUS Homo sapiens FOR II protein mRNA, complete cds.  
DEFINITION  
ACCESSION AF227527  
VERSION AF227527.1 GI:8927390  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2219)  
Ried, K., Finniss, M., Hobson, L., Mangelsdorf, M., Dayan, A.,  
Nancarrow, J.K., Woollett, E., Kremmidiotis, G., Gardner, S.,  
Venter, D., Baker, E. and Richards, R.I.  
Common chromosomal fragile site FRA16D sequence: identification of  
the FOR gene spanning FRA16D and homozygous deletions and  
translocation breakpoints in cancer cells  
Hum. Mol. Genet. 9 (11), 1651-1663 (2000)  
JOURNAL MEDLINE 20320695  
PUBMED 10861292  
REFERENCE 2 (bases 1 to 2219)  
Ried, K., Finniss, M., Hobson, L., Mangelsdorf, M. and Richards, R.I.  
Direct Submission  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
1. .2219  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"



		<div><div>/chromosome="16"</div><div>/map="16q23.2"</div><div>/note="FRA16D"</div><div>105..1349</div><div>/codon_start=1</div><div>/product="FOR II protein"</div><div>/protein_id="AA082054.1"</div><div>/db_xref="GI:8927391"</div><div>/translation="MAALRYAGLDDTDEDELPGWERTTKDGVVYANHTEKTCW EHPKTKRKRKRVAGDLPYGEQETDENGQVFFVDHINKRTTYLDPLRAFTVDDNPFTKT TRGRYDGSYTTAMEILOQRDTGKVVVTGANSIGFTAKSPALHGHVILACRNMAR ASFAVSRILEEWEKHEKVTWLDLALLRSVQHFAEAFKAKNPVPLHVLVNAATFALPD SLTKDGLERTFPQVNHLEGHFVYLQQLDVLKRSAPARVIVSSSESRFTDINDSLGKLD FSRLSATKNDYMWAMLYNRSKLNILFSEHRLRLSPRGVTSNAVHPGNMWSNIHRS WVVTLLFTLARPPTKSMQOGAATTVYCAAVPELEGLGGMVFNCCRCMPSPEAQSEE TARTIWSLSERLIQERLGSQS"</div></div>			
		ORIGIN			
Query Match		97.8%; Score 2213.4; DB 9; Length 2219;			
Best Local Similarity		99.8%; Pred. No. 0;			
Matches 2214; Conservative 2; Mismatches 3; Indels 0; Gaps 0;					
QY	22	TCGGGCCCGACGCGCGGGTCTCGTTTGGAGCGGAGTGAGTTCCTGACGAGTGGAC	81		
DB	1	TCGGGCCCGACGCGCGGGTCTCGTTTGGAGCGGAGTGAGTTCCTGACGAGTGGAC	60		
QY	82	CCGSCAGCGGCGATAGGGGGCGAGTGCCTCCACAGTCAGCCATGGCAGCGCTGGCT	141		
DB	61	CCGSCAGCGGCGATAGGGGGCGAGTGCCTCCACAGTCAGCCATGGCAGCGCTGGCT	120		
QY	142	ACGCGGGCTGGACGACACGACAGTGGAGGACGAGTGCCTCCGGGCTGGGAGGAGAA	201		
DB	121	ACGCGGGCTGGACGACACGACAGTGGAGGACGAGTGCCTCCGGGCTGGGAGGAGAA	180		
QY	202	CCACCAAGGACGCTGGTGTACTACGCCAATCACACCGAGGAGAGACTCAGTGGGAAC	261		
DB	181	CCACCAAGGACGCTGGTGTACTACGCCAATCACACCGAGGAGAGACTCAGTGGGAAC	240		
QY	262	ATCCAAAACCTGAAAAGAAAACGAGTGGCAGGAGATTTGCCATACGATGGGAACAAG	321		
DB	241	ATCCAAAACCTGAAAAGAAAACGAGTGGCAGGAGATTTGCCATACGATGGGAACAAG	300		
QY	322	AAACTGATGAGAACGGCAAGTGTGTTTGTGACCATATAATAAAGAACCACTACT	381		
DB	301	AAACTGATGAGAACGGCAAGTGTGTTTGTGACCATATAATAAAGAACCACTACT	360		
QY	382	TGGACCCAGACTGGCGTTTACTGTGATGATATTCGACCAAGCCAAACCCGGCAAA	441		
DB	361	TGGACCCAGACTGGCGTTTACTGTGATGATATTCGACCAAGCCAAACCCGGCAAA	420		
QY	442	GATACGAGCGCAGCACCACTGCCATGGAATTTCTCAGGGCCGGGATTTCACTGGCAAG	501		
DB	421	GATACGAGCGCAGCACCACTGCCATGGAATTTCTCAGGGCCGGGATTTCACTGGCAAG	480		
QY	502	TGGTTGTGTCACTGGAGCTAAATTCAGGAATPAGGGTTTCAAAACCGCAAGTCTTTGCC	561		
DB	481	TGGTTGTGTGTCACTGGAGCTAAATTCAGGAATPAGGGTTTCAAAACCGCAAGTCTTTGCC	540		
QY	562	TCCATGTGACATGTGATCTTGGCTCGCAGGAACATGGCAAGGGCGAGTGAAGCAGTGT	621		
DB	541	TCCATGTGACATGTGATCTTGGCTCGCAGGAACATGGCAAGGGCGAGTGAAGCAGTGT	600		
QY	622	CACGCATTTTAGAAGATGGCATAAAGCAAGGTAGAAGCAATGACCTCGACCTCGCTC	681		
DB	601	CACGCATTTTAGAAGATGGCATAAAGCAAGGTAGAAGCAATGACCTCGACCTCGCTC	660		
QY	682	TGCTCCGTAGCGTCAGCATTTTGTCTGAAGCATTCAGGGCCAGAAATGTGCTCTTCATG	741		
DB	661	TGCTCCGTAGCGTCAGCATTTTGTCTGAAGCATTCAGGGCCAGAAATGTGCTCTTCATG	720		
QY	742	TGCTTGTGTCAACGACGACAACTTTTGTCTTACCTGGAGTCTCACAAAGATGGCTGG	801		
DB	721	TGCTTGTGTGTCAACGACGACAACTTTTGTCTTACCTGGAGTCTCACAAAGATGGCTGG	780		

QY	802	AGACCACTTTCAAGTGAATCATCTGGGGCACCTTCTACCTTGTCCAGTCTCTCCAGGATG	861
DB	781	AGACCACTTTCAAGTGAATCATCTGGGGCACCTTCTACCTTGTCCAGTCTCTCCAGGATG	840
QY	862	TTTTGTGCCGCTCAGCTCCTGCCCGTGTCAATTGTGGTCTCCTCAGAGTCCCATGATTA	921
DB	841	TTTTGTGCCGCTCAGCTCCTGCCCGTGTCAATTGTGGTCTCCTCAGAGTCCCATGATTA	900
QY	922	CAGATATTAAGCACTCCTTGGGAAAACCTGGAATTCAGTCCGCTCTCTCCACAAAACG	981
DB	901	CAGATATTAAGCACTCCTTGGGAAAACCTGGAATTCAGTCCGCTCTCTCCACAAAACG	960
QY	982	ACTATTGGGCGATCCTGGCTTATAACAGGTCCAAAGCTCTGCAACATCTCTCTCCAAACG	1041
DB	961	ACTATTGGGCGATCCTGGCTTATAACAGGTCCAAAGCTCTGCAACATCTCTCTCCAAACG	1020
QY	1042	AGTCGACCGTGGCTCTCCCCACGCGGGGTACGTGAAACGACAGTGTGATCTCTGGAATA	1101
DB	1021	AGTCGACCGTGGCTCTCCCCACGCGGGGTACGTGAAACGACAGTGTGATCTCTGGAATA	1080
QY	1102	TGATGTACTCCAACTTCATCGCAGCTGGTGGGTGTACACACTGCTGTCTTACCTTGGGA	1161
DB	1081	TGATGTACTCCAACTTCATCGCAGCTGGTGGGTGTACACACTGCTGTCTTACCTTGGGA	1140
QY	1162	GGCCTTTCAACCAAGTCCATGCAACAGGAGCTGCGCACACCGTGTACTGTCTGTCTGCC	1221
DB	1141	GGCCTTTCAACCAAGTCCATGCAACAGGAGCTGCGCACACCGTGTACTGTCTGTCTGCC	1200
QY	1222	CAGAACTGGAGGGTCTGGAGGGATGATCTTCAAACAACCTGCTGCGGTGATGCGCTCAC	1281
DB	1201	CAGAACTGGAGGGTCTGGAGGGATGATCTTCAAACAACCTGCTGCGGTGATGCGCTCAC	1260
QY	1282	CAGAACTGGAGGGTCTGGAGGGATGATCTTCAAACAACCTGCTGCGGTGATGCGCTCAC	1341
DB	1261	CAGAACTGGAGGGTCTGGAGGGATGATCTTCAAACAACCTGCTGCGGTGATGCGCTCAC	1320
QY	1342	AAGAACTGGAGGGTCTGGAGGGATGATCTTCAAACAACCTGCTGCGGTGATGCGCTCAC	1401
DB	1321	AAGAACTGGAGGGTCTGGAGGGATGATCTTCAAACAACCTGCTGCGGTGATGCGCTCAC	1380
QY	1402	CGCCTCTGTGTGTCCTCTCAGCAAGTGCAGGGCTGGGCGCCCTTCAAATGTCCTCC	1461
DB	1381	CGCCTCTGTGTGTCCTCTCAGCAAGTGCAGGGCTGGGCGCCCTTCAAATGTCCTCC	1440
QY	1462	AACACAGATCGCAAGAGTAAAGGAATAGAGCAGTCAACACAGAGTGAAAATCTTAA	1521
DB	1441	AACACAGATCGCAAGAGTAAAGGAATAGAGCAGTCAACACAGAGTGAAAATCTTAA	1500
QY	1522	GTACCAATGGGAAGCAGGGAATCTCTGGGGTAAAGTATCACTTTCTGGGGCTGGGCTAG	1581
DB	1501	GTACCAATGGGAAGCAGGGAATCTCTGGGGTAAAGTATCACTTTCTGGGGCTGGGCTAG	1560
QY	1582	GCAATGCTCTTTGCTTTCTGTGTGGTGGCTGTTGAAAAGTAAAAACCTGCTTGGTGTG	1641
DB	1561	GCAATGCTCTTTGCTTTCTGTGTGGTGGCTGTTGAAAAGTAAAAACCTGCTTGGTGTG	1620
QY	1642	TAGGTTCCGTATCTCCCTGGAGAGCAGCAGCAATCTCTTTCTTTTACTGTTATAGAT	1701
DB	1621	TAGGTTCCGTATCTCCCTGGAGAGCAGCAGCAATCTCTTTCTTTTACTGTTATAGAT	1680
QY	1702	AGCTGAGGTCCCTCTGCTCCATCCAGTACCACACGCGCCACCTGACAGCGGGGCT	1761
DB	1681	AGCTGAGGTCCCTCTGCTCCATCCAGTACCACACGCGCCACCTGACAGCGGGGCT	1740
QY	1762	GGCCTTCTCTACTTAGGGGAAGAAAAGCAAGTGTTCACCTGCTCTCTTGTGCAATGATCC	1821
DB	1741	GGCCTTCTCTACTTAGGGGAAGAAAAGCAAGTGTTCACCTGCTCTCTTGTGCAATGATCC	1800
QY	1822	AGGAGATAATTTGTTTCAATTCATCTGACCAAGACTGAGCCAGCTTAGCACTGCTGGGA	1881
DB	1801	AGGAGATAATTTGTTTCAATTCATCTGACCAAGACTGAGCCAGCTTAGCACTGCTGGGA	1860

```
QY 1862 GACAAATCTGAGACCTTGTCCAGCAGTGGAGATGACAGTGAACACCCAGAGGAGTAG 1941
Db 1861 GACAAATCTGAGACCTTGTCCAGCAGTGGAGATGACAGTGAACACCCAGAGGAGTAG 1920
QY 1942 AATACGAGAACTACAGGTCGCAAGTACTTGTCTATAGACTCCCTTGTCTAATGCTATGC 2001
Db 1921 AATACGAGAACTACAGGTCGCAAGTACTTGTCTATAGACTCCCTTGTCTAATGCTATGC 1980
QY 2002 AAAAAATCTTTAGAGATTAATAACAAATTTTCAAAATCATTTCTTAGATACCTTGAAAGG 2061
Db 1981 AAAAAATCTTTAGAGATTAATAACAAATTTTCAAAATCATTTCTTAGATACCTTGAAAGG 2040
QY 2062 CAGGAAGGAGCGGTATATATTAAGAATAACACAGATATTTTGGGGGCGAGAGATAAA 2121
Db 2041 CAGGAAGGAGCGGTATATATTAAGAATAACACAGATATTTTGGGGGCGAGAGATAAA 2100
QY 2122 ACGTTAGTTAATCCCTTGTCTGCTCAATACAGTCTCAGTCTCTTGTCTTACATTTGTA 2181
Db 2101 ACGTTAGTTAATCCCTTGTCTGCTCAATACAGTCTCAGTCTCTTGTCTTACATTTGTA 2160
QY 2182 CTTAAACCTCTGCTGTGCTCGCATCTCTATGCTTTAATAAAGAACATGCTTGAATATC 2240
Db 2161 CTTAAACCTCTGCTGTGCTCGCATCTCTATGCTTTAATAAAGAACATGCTTGAATATC 2219

RESULT 7
HSU13395
LOCUS HSU13395 1475 bp mRNA linear PRI 15-SEP-1994
DEFINITION Human oxidoreductase (HHCMA56) mRNA, complete cds.
ACCESSION U13395
VERSION U13395.1 GI:538131
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 1475)
Gmerek,R.E. and Medford,J.I.
The complete sequence of a human hippocampus gene (HHCMA56) shows
homology to developmental genes from Arabidopsis and Brassica napus
Unpublished
2 (bases 1 to 1475)
Gmerek,R.E.
Direct Submission
Submitted (11-AUG-1994) Ronald E. Gmerek, Biology, Eberly College
of Science, The Pennsylvania State University, 506 Wartik
Laboratory, University Park, PA 16802, USA
Location/Qualifiers
1..1475
/organism="Homo sapiens"
/mol_type="mRNA"
/isolate="two year old female"
/db_xref="taxon:9606"
/chromosome="16"
/sex="female"
/tissue_type="hippocampus"
/clone_lib="hippocampus library, Stratagene catalog number
936205"
/dev_stage="juvenile"
/note="DSEG number: D168432E"

1..1475
/gene="HHCMA56"
19..1131
/gene="HHCMA56"
/codon_start=1
/product="oxidoreductase"
/protein_id="AAA21465.1"
/db_xref="GI:538132"

translation="MARSEAVSRILEEWHKAKVEAMTLDIALRSVQHFAEFKAKN
VPLHVCNAAATFALPWSLTGDLTTFQVNLGHFVLVOLLPGMFCQAQLLPVSLWS
PQSDIQLTTPWNTSVASLOKTTIGRCWLTITGPSSATSSPTSCVSPTRGH
VERSDRWKMDVLQCHVHTAVYLGEAFHQVHATGSCHEHVLCCCPRTGSSRD
VLQQLPLHALTRSSRRDGPVGLSERLIQERLAASPAKWSSERMGTHTRPVCVPS
```

```
ORIGIN
Query Match 60.6%; Score 1372.4; DB 9; Length 1475;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1459; Conservative 0; Mismatches 11; Indels 10; Gaps 7;

QY 579 ATCTTGGCTCAGGAACATCGCAAGCGGAGTGAAGCAGTGTACGCAATTTTAGAAGAA 638
Db 1 ATCTTGGCTCAGGAACATCGCAAGCGGAGTGAAGCAGTGTACGCAATTTTAGAAGAA 60
QY 639 TGGCATAAAGCAAGGTAGAAGCAATGACCCCTGGACCTCGCTCTGCTCCGTAGCGTGCAG 698
Db 61 TGGCATAAAGCAAGGTAGAAGCAATGACCCCTGGACCTCGCTCTGCTCCGTAGCGTGCAG 120
QY 699 CATTTTGTGAAGCATTTCAAGGCCAAGATGTGCTCTTTCATGTGCTTGTGTGCAACGCA 758
Db 121 CATTTTGTGAAGCATTTCAAGGCCAAGATGTGCTCTTTCATGTGCTTGTGTGCAACGCA 180
QY 759 GCAACTTTTGTCTACCCCTGGAGTCTACCAAGATGGCTGGAGACCAACCTTTCAAAGTG 818
Db 181 GCAACTTTTGTCTACCCCTGGAGTCTACCAAGATGGCTGGAGACCAACCTTTCAAAGTG 240
QY 819 AATCATCTGGGGCACTTCTACCTTGTCCAGTCTCTCC--AGGATGTTTGTGCCGCTCAG 876
Db 241 AATCATCTGGGGCACTTCTACCTTGTCCAGTCTCTCC--AGGATGTTTGTGCCGCTCAG 300
QY 877 CTCCTGCCCGTGCATTTGTGTCTCCTCAGATGCCATCGATTTACAGATATTACGACT 936
Db 301 CTCCTGCCCGTGCATTTGTGTCTCCTCAGATGCCATCGATTTACAGATATTACGACT 360
QY 937 CTTTGGGAAAACTGACCTTCAGTGCCTCTCTCCAAACAAAAACGACTATTGGGCGATGC 996
Db 361 CTTTGGGAAAACTGACCTTCAGTGCCTCTCTCCAAACAAAAACGACTATTGGGCGATGC 420
QY 997 TGGCTTTAAACAGTTCAGCTCTGCAACATCTCTTCTCCAAACGAGTGCACCGTCCG 1056
Db 421 TGGCTTTAAACAGTTCAGCTCTGCAACATCTCTTCTCCAAACGAGTGCACCGTCCG 480
QY 1057 TCTCCCAACGCGGGTTCAGTGCAGACGAGTG--CATCTGGAATATGATGACTCCA 1113
Db 481 TCT--CCCAACGCGGGTTCAGTGCAGACGAGTG--CATCTGGAATATGATGACTCCA 539
QY 1114 ACATTCATCGCAGCTGGTGGGTGTACACACTGCTGTTTACCTTGGCGAGGCTTTTCA 1173
Db 540 ACATTCATCGCAGCTGGTGGGTGTACACACTGCTGTTTACCTTGGCGAGGCTTTTCA 599
QY 1174 AGTCCATGCAACAGGAGCTGCCACCCCGTGTACTGTGCTGTCTCCAGAACTGGAGG 1233
Db 600 AGTCCATGCAACAGGAGCTGCCACCCCGTGTACTGTGCTGTCTCCAGAACTGGAGG 659
QY 1234 GTCTGGGAGGATGTACTTCAACAACTGCTGCCGCTCATGCCCTCACCAGAACTCAGA 1293
Db 560 GTCTAGAGGATGTACTTCAACAACTGCTGCCGCTCATGCCCTCACCAGAACTCAGA 719
QY 1294 GCGAAGAGACGCGCCCGGACCTGTGGGCGCTCAGCGAGGCTGATCCAGAAAGCGTTG 1353
Db 720 GCGAAGAGACGCGCCCGGACCTGTGGGCGCTCAGCGAGGCTGATCCAGAAAGCGTTG 777
QY 1354 GCAGCCAGTCCGGCTAAGTGAGCTCAGAGCGGATGGGCAACACCCGCCCTGTGTGT 1413
Db 778 GCAGCCAGTCCGGCTAAGTGAGCTCAGAGCGGATGGGCAACACCCGCCCTGTGTGT 837
QY 1414 GTCCCTTCACGCAAGTGCCAGGGCTGGGCGCTTCCAAATGTCCCTCCACACAGATCCG 1473
Db 838 GTCCCTTCACGCAAGTGCCAGGGCTGGGCGCTTCCAAATGTCCCTCCACACAGATCCG 896
QY 1474 CAAGAGTAAAGGAAATAAGAGCAGTCAACACAGAGTGAATAAATCTTAAGTACCAATGGGA 1533
Db 897 CAAGAGTAAAGGAAATAAGAGCAGTCAACACAGAGTGAATAAATCTTAAGTACCAATGGGA 956
QY 1534 AGCAGGAAATTCCTGGGGTAAAGTATCACTTTTCTGGGGCTGGGCTAGGCATAGGTCTCT 1593
```

RKCAGPLENPVPTQIRKSKGNKSTHNRKNLKYQWEAGNSWGKVSLEFWGWARHSLC  
PLVACLKVKTWLACRFRISLEKHQFSSFYCIA"

111

957 AGCAGGGAATTCCTGGGGTAAAGATACATCTTTCTGGGGCTAGGATAGGTCCT 1016

1594 TTGCTTTCTGGTGGTGGCCCTGTTTGAAGTAAACCTCTGGTGTGTTAGGTCCTGAT 1653

1017 TTGCTTTCTGGTGGTGGCCCTGTTTGAAGTAAACCTCTGGTGTGTTAGGTCCTGAT 1076

1654 CTCCTGGGAAGCACCAGCAATCTCTTTCTTTTACTGTATAGATAGCCTGAGTCC 1713

1077 CTCCTGGGAAGCACCAGCAATCTCTTTCTTTTACTGTATAGATAGCCTGAGTCC 1136

1714 CTCCTGGTCCATCCAGTACACCAAGCCACCTATGTCAGCGGGGCTGCCCTTCTCTA 1773

1137 CCTCGT-CCATCCAGTACACCAAGCCACCTATGTCAGCGGGGCTGCCCTTCTCTA 1195

1774 CTTAGGGAAGAAAAGCAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1833

1196 CTTAGGGAAGAAAAGCAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1255

1834 TTTCTATTCT 1893

1256 TTTCTATTCT 1315

1894 AACCTTGTCCAGCCAGTGGAGTACAGTGACACCCAGAGGGAGTATAGATACGAGAAC 1953

1316 AACCTTGTCCAGCCAGTGGAGTACAGTGACACCCAGAGGGAGTATAGATACGAGAAC 1375

1954 TACAGGTGCAAGTCTGTATAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2013

1376 TACAGGTGCAAGTCTGTATAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1435

2014 AGAGATTATACAAATTTTCAATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2053

1436 AGAGATTATACAAATTTTCAATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1475

RESULT 8

AF395123

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. 1732

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

134. .838

/note="aberrant WWOX transcript found in cancer cells"

/codon\_start=1

/evidence=not\_experimental

/product="WWOXdelta6-8"

/protein\_id="AAK81727.1"

/db\_xref="GI:15028463"

CDS

/translation="MAALRYAGLDLTDSELPFGWEERTTKDGMVYVYAHNTEKTDW  
EHPKTKRVRAGDLPVGEQETDENGQVFFVDHINKRTYLDPLRAFTVDDNPKPT  
TRQVYDSTGTFAMEILQDRDFTGKVVVTVGANSIGFETAKSFAHGHVILAENMAR  
ASEAVRILIEWOQGAATTVYCAAVPELBLEGLGMFYNNCCRCMPSPQAQSEETARTLW  
AUSERLIQRLRSGSQS"

ORIGIN

Query Match 51.9%; Score 1174; DB 9; Length 1732;  
Best Local Similarity 76.1%; Pred. No. 2.1e-310;  
Matches 1724; Conservative 0; Mismatches 0; Indels 540; Gaps 1;

QY 1 GCAGTGGCAGGCGTGGAGCGTGGGCCCCGACGCGCGCGGGGTCTCGTTTGGAGCGGAG 60

DB 9 GCAGTGGCAGGCGTGGAGCGTGGGCCCCGACGCGCGGGGTCTCGTTTGGAGCGGAG 68

QY 61 TGAGTTCTCAGCGAGTGGACCCGCGCAGCGGGCGATAGGGGGGCGAGTGCCTCCAGT 120

DB 69 TGAGTTCTCAGCGAGTGGACCCGCGCAGCGGGCGATAGGGGGGCGAGTGCCTCCAGT 128

QY 121 CAGCCATGGCAGCGCTGCGCTACGCGGGGTGGACACGACACGACAGTGGAGCAGTGC 180

DB 129 CAGCCATGGCAGCGCTGCGCTACGCGGGGTGGACACGACACGACAGTGGAGCAGTGC 188

QY 181 CTCGGGCTGGGAGGAGAACCCACCAAGGACCGCTGGGTTTACTACGCCAATCACACCG 240

DB 189 CTCGGGCTGGGAGGAGAACCCACCAAGGACCGCTGGGTTTACTACGCCAATCACACCG 248

QY 241 AGAGAAAGTCTAGTGGGAAATCCAAAATCTGAAAAAAGAAAAACGAGTGGCAGGAGAT 300

DB 249 AGAGAAAGTCTAGTGGGAAATCCAAAATCTGAAAAAAGAAAAACGAGTGGCAGGAGAT 308

QY 301 TGCATACGATGGGAAACAAAGAACTGATGAGAACCGACAAAGTGTCTTTTGTGACCAT 360

DB 309 TGCATACGATGGGAAACAAAGAACTGATGAGAACCGACAAAGTGTCTTTTGTGACCAT 368

QY 361 TAAATAAAGAACCACTACTTGGACCAAGCTGGCGTTTACTGTGGATGATTAATCCGA 420

DB 369 TAAATAAAGAACCACTACTTGGACCAAGCTGGCGTTTACTGTGGATGATTAATCCGA 428

QY 421 CCAAGCCACCAACCGGCAAGATACGACGCGGACCACTGCGCATGGAATCTCCAGG 480

DB 429 CCAAGCCACCAACCGGCAAGATACGACGCGGACCACTGCGCATGGAATCTCCAGG 488

QY 481 GCGGGATTTCTGCAAAAGTGTGTGTCTACTGGAGCTAAATTCAGGAATAGGGTTCG 540

DB 489 GCGGGATTTCTGCAAAAGTGTGTGTCTACTGGAGCTAAATTCAGGAATAGGGTTCG 548

QY 541 AAACCCCAAGTCTTTTGGCCCTCCATGGTGCACATGTGATCTTGGCTGCGAGAACATGG 600

DB 549 AAACCCCAAGTCTTTTGGCCCTCCATGGTGCACATGTGATCTTGGCTGCGAGAACATGG 608

QY 601 CRAAGGGGAGTGAAGCAGTGTCAAGCATTTTGAAGAAATGCGCATAAAGCCAGTAGAAG 660

DB 609 CRAAGGGGAGTGAAGCAGTGTCAAGCATTTTGAAGAAATGCGCATTTTGAAGAAATG 648

QY 661 CAATGACCCCTGGACCTGCTCTCGTAGCGTGGAGCATTTTGTGTAAGCATTTCAAGG 720

DB 649 CAATGACCCCTGGACCTGCTCTCGTAGCGTGGAGCATTTTGTGTAAGCATTTCAAGG 648

QY 721 CCAAGAAATGTCCTCTTCAATGTCTGTGTGCAACGACGCAATTTTGTCTTACCTCGGA 780

DB 649 CCAAGAAATGTCCTCTTCAATGTCTGTGTGCAACGACGCAATTTTGTCTTACCTCGGA 648

QY 781 GTCTCACCAGAGTGGCTGGAGACCACTTTTCAAGTGAATCATCTGGGCACTTCTTACC 840

DB 649 GTCTCACCAGAGTGGCTGGAGACCACTTTTCAAGTGAATCATCTGGGCACTTCTTACC 648

QY 841 TTGTCCAGTCTCTCCAGGAGTTTGTGCGCTCAGCTCTCGCGGTGTCATTTGTGTCT 900

DB 649 TTGTCCAGTCTCTCCAGGAGTTTGTGCGCTCAGCTCTCGCGGTGTCATTTGTGTCT 648

QY 901 CCTCAGAGTCCCATCGATTTTACAGATTTTAAAGCATCTCTTGGGAAACCTGGACTTCAGTC 960



101 GGCGAGGTGCTCCACAGTCAGCCATGGCAGCGCTGCTAGCGGGGCTGGACGAC 160  
102 |||||  
61 GGCGAGGTGCTCCACAGTCAGCCATGGCAGCGCTGCTAGCGGGGCTGGACGAC 120  
103 |||||  
161 GGACAGTGGAGCGAGCTGCTCGGGCTGGAGGAGAGAACACCAAGACCGCTGGT 220  
104 |||||  
121 GGACAGTGGAGCGAGCTGCTCGGGCTGGAGGAGAGAACACCAAGACCGCTGGT 180  
105 |||||  
221 TTACTACGCCAATCACACGAGGAGAGACTCAGTGGGAATCCAAAATCTGAAAAAG 280  
106 |||||  
181 TTACTACGCCAATCACACGAGGAGAGACTCAGTGGGAATCCAAAATCTGAAAAAG 240  
107 |||||  
281 AAAACAGTGGAGGAGATTTGCCATACGATGGAAACAAAGAACTGATGAGAACGACA 340  
108 |||||  
241 AAAACAGTGGAGGAGATTTGCCATACGATGGAAACAAAGAACTGATGAGAACGACA 300  
109 |||||  
341 AGTGTGTTTTTTGTTGACCATATAATAAAGAACCACTCTTGACCAAGACTGGCGTT 400  
110 |||||  
301 AGTGTGTTTTTTGTTGACCATATAATAAAGAACCACTCTTGACCAAGACTGGCGTT 360  
111 |||||  
401 TACTGTGGATGATTAATCCGACCAAGCAACACCCGCAAGATACAGCGCAGCACAC 460  
112 |||||  
361 TACTGTGGATGATTAATCCGACCAAGCAACACCCGCAAGATACAGCGCAGCACAC 420  
113 |||||  
461 TGCATGGAAATTTCTCCAGGCGCGGATTTCACTGGCAAGTGTGTGTCTCACTGGAGC 520  
114 |||||  
421 TGCATGGAAATTTCTCCAGGCGCGGATTTCACTGGCAAGTGTGTGTCTCACTGGAGC 480  
115 |||||  
521 TAATTCAGGAATAGGTTTCAAAACCGCAAGTCTTTTGGCCCTCCATGGTGCACATGTGAT 580  
116 |||||  
481 TAATTCAGGAATAGGTTTCAAAACCGCAAGTCTTTTGGCCCTCCATGGTGCACATGTGAT 540  
117 |||||  
581 CTTGGCTGCAGGAATGTCGACGAGGCGAGTGAAGTGTCTGCGATTTTAGAAGATG 640  
118 |||||  
541 CTTGGCTGCAGGAATGTCGACGAGGCGAGTGAAGTGTCTGCGATTTTAGAAGATG 600  
119 |||||  
641 GCATAAGCCAAAGTGAAGCAATGACCTGGACCTCGCTCTGCTCGTGAAGTGGAGCA 700  
120 |||||  
601 GCATAAGCCAAAGTGAAGCAATGACCTGGACCTCGCTCTGCTCGTGAAGTGGAGCA 660  
121 |||||  
701 TTTTGTGTGAAGCAATCAAGCCCAAGATGTCCTTTTCATGTGTTGTGCAACGAGC 760  
122 |||||  
661 TTTTGTGTGAAGCAATCAAGCCCAAGATGTCCTTTTCATGTGTTGTGCAACGAGC 720  
123 |||||  
761 AACTTTTGTCTACCTGGAGTCTCAACCAAGATGGCTGGAGACCACTTTCAAGTGAA 820  
124 |||||  
721 AACTTTTGTCTACCTGGAGTCTCAACCAAGATGGCTGGAGACCACTTTCAAGTGAA 780  
125 |||||  
821 TCATCTGGGCACTTCTACCTTGTCCAGCTCCTCCAGGATGTTTGTGCGGCTCAGCTCC 880  
126 |||||  
781 TCATCTGGGCACTTCTACCTTGTCCAGCTCCTCCAGGATGTTTGTGCGGCTCAGCTCC 840  
127 |||||  
881 TGCCCGGTGTCATTTGCTCTCCAGAGTCCCAATGATTTACAGATATTAACAGCTCTT 940  
128 |||||  
841 TGCCCGGTGTCATTTGCTCTCCAGAGTCCCAATGATTTACAGATATTAACAGCTCTT 900  
129 |||||  
941 GGGAAACTGGACTTCACTGCTCTCTCCAAACAAACCACTATTTGGCGATGCTGGC 1000  
130 |||||  
901 GGGAAACTGGACTTCACTGCTCTCTCCAAACAAACCACTATTTGGCGATGCTGGC 960  
131 |||||  
1001 TTATAACAGGTCCAAGTCTGCAACATCTCTTCTCCAAACAGCTGACCGTGCCTCTC 1060  
132 |||||  
961 TTATAACAGGTCCAAGTCTGCAACATCTCTTCTCCAAACAGCTGACCGTGCCTCTC 1020  
133 |||||  
1061 CCCACGCGGGTCACTGTCGAACGAGTGCATCTCTGGAATATGATGCTACTCCAACTTCA 1120  
134 |||||  
1021 CCCACGCGGGTCACTGTCGAACGAGTGCATCTCTGGAATATGATGCTACTCCAACTTCA 1080  
135 |||||  
1121 TCCGAGCTGGTGGGTGATACACATGCTGTTTACCTTGGCGAGGCGCTTTCACCAAGTCCAT 1180  
136 |||||  
1081 TCCGAGCTGGTGGGTGATACACATGCTGTTTACCTTGGCGAGGCGCTTTCACCAAGTCCAT 1140  
137 |||||

QY 1181 G 1181  
Db 1141 G 1141

RESULT 10  
BC014716  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BC014716 2269 bp mRNA linear ROD 04-OCT-2003  
Mus musculus WW domain-containing oxidoreductase, mRNA (cDNA clone  
MGC:25975 IMAGE:4241066), complete cds.  
BC014716  
BC014716.1 GI:15928475  
MGC.  
Mus musculus (house mouse)

REFERENCE  
AUTHORS

1 (bases 1 to 2269)  
Klausner, R.D., Collins, F.S., Wagner, L., Shennen, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Scheetz, T.E., Brownstein, M.J., Udell, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullah, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,  
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E.,  
Schmerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences

TITLE

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

2 (bases 1 to 2269)  
Strausberg, R.  
Direct Submission  
Submitted (01-OCT-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

REMARK  
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Lounseged, H.,  
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,  
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 31 Row: f Column: 1  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 9625011.

FEATURES

source

1. .2269  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"

QY	587	CTGCAGGAACATGGCAAGGCGGAGTGAAGCAGTGTCAAGCATTTTATAGAAATGGCATAA	646
Db	538	CTGTAGGAACCTGAGCAGAGCAGCGAGGCGGTGTCCGCGATCCTGGAGAGTGGCACAA	597
QY	647	AGCCAAAGGTAGAAGCAATGACCCCTGGAACCTCGCTCTGCTCCGTAGCGTGCAGCATTTTGC	706
Db	598	AGCCAAAGGTGGAAGCAATGACCCCTGGAACCTGGCCGCTGCTCCGTAGCGTGCAGCATTTTGC	657
QY	707	TGAAGCATTTCAAGGCCCAAGATGTGCCTCTTCAATGTGCTTGTGCAACCCAGCAACTTT	766
Db	658	CGAAGCATTTCAAGGCCCAAGATGTGCTCTTTCAAGTGTCTTGTGTATATGACGGACGTT	717
QY	767	TGCTCTACCCCTGGAGTCTCAACAAAGATGCCCTGGAGACCACTTTTCAAGTGAATCATCT	826
Db	718	TGCTCTGCCCTGGGCGCTCAACGAAGGATGGCTGGAGACAAACCTTTTCAAGTGAACCACT	777
QY	827	GGGGCACTTCTACCTTGTCCAGCTCTCTCCAGATATTTTGTGTGCGCGCTCAGCTCCTGCCCG	886
Db	778	GGGACATTTCTACCTTGTCCAACTCCTCGAGGATGTCTGTGTGCGCTCATCCCGCGCGCG	837
QY	887	TGTCATTTGTGCTCTCCTCAGAGTCCCATCGATTTACAGATATTTAAACGACTCCTTGGGAAA	946
Db	838	TGTCATCTGTGCTCTCTTCGAGTCCCAAGATTTACAGATATTTAATGACTCTCCCGGAAA	897
QY	947	ACTGGAATTTCAAGTCGCTCTCTCCAAACAAAAACGACTATTGGCGGATGTGGCTTATAA	1006
Db	898	ACTTGACCTCAGCGCGCTGTCTCCAACGAGGAGGACTACTGGGCCATGTGGCTTATAA	957
QY	1007	CAGGTCGAAGCTCTGCAACATCTCTTTTCCAAAGAGCTGCAAGTGTGCTCTCCCGCAG	1066
Db	958	CAGATCTAAACTCTGCAATATCTCTTCTCCAAATGAATCGACCGTGGCTCTCCCGCGC	1017
QY	1067	CGGGGTCAAGTCGCAACGCGAGTGCATCTCGGAATATGATGTAATCCAAATTCATCGCAG	1126
Db	1018	CGGGGTCAATCAACGCGAGTGCACCTTGGAAACATGATGTAATCGACCATTCATCGCAA	1077
QY	1127	CTGGTGGGTGTACACACTGCTGTTTACCTTGGCGAGGCGCTTTCAACAAATGCCATGCAACA	1186
Db	1078	CTCATGGGTGTACAGTGTCTGTTTCACTTGGCGAGGCGCTTCAACAAATGCCATGCAACA	1137
QY	1187	GGGAGCTGCCACCAACCGTGTACTGTGCTGTCCAGAACTGGAGGGTCTGGAGGGAT	1246
Db	1138	GGGAGCTGTCTACACTGTCTACTGTGAGTGTGCGGCCCAAGAGCTGGAGGGAT	1197
QY	1247	GTACTTTCAACAACTGTCTGGCTGTGATGCCCTCACACAGAGCTCAGAGCGAAGAGAGCGC	1306
Db	1198	GTATTTTCAACAACTGTCTGGCTGTGCTGTCCATCCGAGGAGGCTCAGAGTGAAGAGACGC	1257
QY	1307	CCGGAACCTGTGGGCGCTCAGCGAGAGGCTGATCCAAAGACGGCTTGGCAGCCAGTCCGG	1366
Db	1258	TCCGGCCCTGTGGAACTCAGTGAAGAGGCTCATCAGACCGAGTGGTAGTCCATCCAG	1317
QY	1367	CTAAGTGGAGCTCAGAGCGGATGGGCACACACACCGCCCTGTGTGTGCCCTCAGCA	1426
Db	1318	CTAATTTGGAACCCATAATTTGGGCAAAACCCACCC-----ATGCTCTGACCT	1368
QY	1427	AGTGCCAGGGCTGGGCGCTTCCAAATGTCCCTCCAAACACAGATCCGACAGATGAAGGA	1486
Db	1369	CATCACAGTCAAGCCCATCATCTCGGGTGCCTTGAACTCCAGTACAGGGCAATGAGAGTA	1428
QY	1487	AATAAGCAGAGTCACAAACAGAGTGAAATCTTTAAGTACCAATGGGAAGCAGGGAAATTC	1546
Db	1429	GATTAAGAACATC-----AATGGGAGAGGTTATACTAGTGGGAAGCAGGGAAATTC	1481
QY	1547	TGGGGTAAAGTATCACTTTTCTGGGCTGGGCTAGGCATAGGCTCTCTTGC-TTTCCTGGT	1605
Db	1482	AAGAG-ATGATGTCCCTTTTCTAGGACTGGTGTAGTGTGGGACTGCTTGTCTTTCTGGT	1540
QY	1606	GGTGGCTGTTGAAGTAAAGTAAACCTGTCTGGTGTGTAGTTCGGTATCTCCCTCGAGAA	1665
Db	1541	GTTGGCCGCTTGAAATGGAATCATGTGGTAGCATAGGTTCCTTGTCTCAATTGTAAAA	1600





Query Match.	48.3%;	Score 1092.8;	DB 9;	Length 1669;
Best Local Similarity	75.1%;	Pred. No.4e-288;		
Matches 1662;	Conservative 0;	Mismatches 7;	Indels 545;	Gaps 2;
QY	47	GTTTGGAGCGGAGTGTGATTTCTCAGACGAGTGGACCCGGCAGCGGGCGATAGGGGGGCCA	106	
DB	1	GTTTGGAGCGGAGTGTGATTTCTCAGCGAGTGGACCCGGCAGCGGGCGATAGGGGGGCCA	60	
QY	107	GGTGCTTCCACAGTACAGCCATGGCAGCGCTGCGCTACGCGGGGCTGACACGACCGACAG	166	
DB	61	GGTGCTTCCACAGTGTAGCCATGGCAGCGCTGCGCTACGCGGGGCTGACACGACCGACAG	120	
QY	167	TGAGGACGAGCTGCTCCGGCTCGGAGGAGAGAACCAACGACCGCTGGGTTTTACTA	226	
DB	121	TGAGGACGAGCTGCTCCGGCTCGGAGGAGAGAACCAACGACCGCTGGGTTTTACTA	180	
QY	227	CGCCAAATCACACCGAGGAGAGACTCAGTGGGAACATCCAAAACTGGAAAAAGAAAACG	286	
DB	181	CGCCAAATCACACCGAGGAGAGACTCAGTGGGAACATCCAAAACTGGAAAAAGAAAACG	240	
QY	287	AGTGGCAGGAGATTGGCATACGATGGGAACAAGAAACTGTATGAGAACCGACAAGTGT	346	
DB	241	AGTGGCAGGAGATTGGCATACGATGGGAACAAGAAACTGTATGAGAACCGACAAGTGT	300	
QY	347	TTTTTGTTCACATATAATAAAGAACCACTACTTGGACCCAGACTGGCGTTTACTGT	406	
DB	301	TTTTTGTTCACATATAATAAAGAACCACTACTTGGACCCAGACTGGCGTTTACTGT	360	
QY	407	GGATGATAATCCGACCAAGCAACCAACCCGGCAAGATACGACGGCAGCACTGCCAT	466	
DB	361	GGATGATAATCCGACCAAGCAACCAACCCGGCAAGATACGACGGCAGCACTGCCAT	420	
QY	467	GGAAATTTCTCAGGCGCGGATTTTCACTGGCAAGTGGTGTGTCTCATGTGAGCTAATTC	526	
DB	421	GGAAATTTCTCAGGCGCGGATTTTCACTGGCAAGTGGTGTGTCTCATGTGAGCTAATTC	480	
QY	527	AGGAATAGGGTTCGAAACCGCAAGTCTTTTGCCCTCCATGGTGCAATGTGATCTTGGC	586	
DB	481	AGGAATAGGGTTCGAAACCGCAAGTCTTTTGCCCTCCATGGTGCAATGTGATCTTGGC	540	
QY	587	CTGCAGAAATGGCAAGGGCGAGTGAAGCAGTGTCCGCAATTTTGAAGATGGCATAA	646	
DB	541	CTGCAGAAATGGCAAGGGCGAGTGAAGCAGTGTCCGCAATTTTGAAGATGGCATAA	594	
QY	647	AGCCAAGGTAGAAGCAATGACCCCTGGACCTCGCTCTGCTCCGTAGCGTGACGCAATTTGC	706	
DB	595	-----	594	
QY	707	TGAAGCATTCAGGCGCAAGAAATGTGCCTCTTCATGTGCTTGTGTGCAACGACGCACTTT	766	
DB	595	-----	594	
QY	767	TGCTCTACCTTGGAGTCTTACCAAGATGGCTGGAGACCACCTTTTCAAGTGAATCATCT	826	
DB	595	-----	594	
QY	827	GGGGCACTTCTACGTTTCCAGCTCTCCAGGATGTTTTTGTGCGGCTCAGCTCTCGCCCG	886	
DB	595	-----	594	
QY	887	TGTCATTTGTTCTCTCAGAGTCCCATCGATTTACAGATATTAAAGACTCTCTTGGGAAA	946	
DB	595	-----	594	
QY	947	ACTGGACTTCAGTCGCTCTCTCCACAAAAACGACTATTGGGCGATGCTGGCTTATAA	1006	
DB	595	-----	594	
QY	1007	CAGGTCGAAGCTCTGCAACATCTCTTCTCCAACGAGCTGACGCTCGCTCTCCCAACG	1066	
DB	595	-----	594	

QY	1067	CGGGCTCAGCTGCGAAGCAGTGCATCTCTGGAAATATGATGTATCTCCAAACATTCATCCGAG	1126
Db	595	-----	594
QY	1127	CTGCTGGGTGTACACACTGCTGTTTACCTTTGGGAGGSCCTTTACCAAGTCCATGCAACA	1186
Db	595	-----GCAACA	600
QY	1187	GGGAGCTGCCACCAACCGTGTACTGTGCTGTCTGCCAGAACTGGAGGGTCTGGAGGGAT	1246
Db	601	GGGAGCTGCCACACACCGTGTACTGTGCTGTGCCAGAACTGGAGGGTCTGGAGGGAT	660
QY	1247	GTACTTCAACAACTGCTGCCGCTGCATGCCCTCACACGAAAGCTCAGAGCGAAGAGCGGC	1306
Db	661	GTACTTCAACAACTGCTGCCGCTGCATGCCCTCACCAAGAGCTCAGAGCGAAGAGAGCGGC	720
QY	1307	CCGAGCCCTCTGGGCGCTCAGCAGAGGCTGATCCCAAGACGCTTGGCAGCGATCCCG	1366
Db	721	CCGAGCCCTCTGGGCGCTCAGCAGAGGCTGATCCCAAGAACGGCTTGGCAGCGATCCCG	780
QY	1367	CTAAGTGGAGCTCAGAGCGGATGGGCACACACCCCGCCCTGTGTGTGCCCTCAGCGA	1426
Db	781	CTAAGTGGAGCTCAGAGCGGATGGGCACACACCCCGCCCTGTGTGTGCCCTCAGCGA	840
QY	1427	AGTGCCAGGGCTGGGCGCCCTTCCAAATGTCCCTCCAAACAGATCCGCAAGATAAAGGA	1486
Db	841	AGTGCCAGGGCTGGGCGCCCTTCCAAATGTCCCTCCAAACAGATCCGCAAGATAAAGGA	900
QY	1487	AATAAGAGCAGTCACACAGAGTGAAATCTTAAGTACCAATGGGACGAGGAAATTC	1546
Db	901	AATAAGAGCATTACACACAGAGTGAAATCTTAAGTACCAATGGGAGCAGGGAATTC	960
QY	1547	TGGGCTAAAGTATCACTTTTCTGGGCTGGGCTAGGCATAGGTCCTCTTTGCTTTCTGGTG	1606
Db	961	TGGGCTAAAGTATCACTTTTCG-----GGGCTAGCATAGGTCCTCTTTGCTTTCTGGTG	1015
QY	1607	GTGGCTGTTTGAAGTAAACAACTGCTGGTGTGTAGGTTCCGTAATCCCTCGGAGAAG	1666
Db	1016	GTGGCTGTTTGAAGTAAACAACTGTTGGCGTGTAGGTTCCGTAATCCCTCGGAGAAG	1075
QY	1667	CACGAGAAATCTCTTTTACTGTTATAGATAGCCTGAGGTCCTCGTCCCTCGTCCATCC	1726
Db	1076	CACGAGAAATCTCTTTTACTGTTATAGATAGCCTGAGGTCCTCGTCCCTCGTCCATCC	1135
QY	1727	AGCTACCAACACGGCCACCACTCGACCGGGGCTGGCCCTCTCTTACTTTAGGGAAGAA	1786
Db	1136	AGCTACCAACACGGCCACCACTCGACCGGGGCTGGCCCTCTCTTACTTTAGGGAAGAA	1195
QY	1787	AAGCAAGTGTTCACTGCTCTCTGCTGCATGTATCCAGGAGATAATTTTCAATCATCCT	1846
Db	1196	AAGCAAGTGTTCACTGCTCTCTGCTGCATGTATCCAGGAGATAATTTTCAATCATCCT	1255
QY	1847	GACCAAGCTGAGCCAGCTTAGCAACTGCTGGGGAGACAAATCTCAGAACCTTGTCCCG	1906
Db	1256	GACCAAGCTGAGCCAGCTTAGCAACTGCTGGGGAGACAAATCTCAGAACCTTGTCCCG	1315
QY	1907	CCAGTGAGGATGACAGTGACACCCAGAGGAGTAGAATACGCAAGAACTACCAAGTGGCAA	1966
Db	1316	CCAGTGAGGATGACAGTGACACCCAGAGGAGTAGAATACGCAAGAACTACCAAGTGGCAA	1375
QY	1967	AGTACTGTGCATAGACTCCTTTGCTTAATGCTATGCAAAAAATCTTTTAGAGATTAAACA	2026
Db	1376	AGTACTGTGCATAGACTCCTTTGCTTAATGCTATACAAAAATCTTTTAGAGATTAAACA	1435
QY	2027	AATTTTCAAAATCAATTCCTTAGATACCTTGAAAGGAGGAAAGCGGTATATATCTTAA	2086
Db	1436	AATTTTCAAAATCAATTCCTTAGATACCTTGAAAGGAGGAAAGCGGTATATATCTTAA	1495
QY	2087	GAATACACAGGATATTTTGGGGCGAGATATAAACGTTAGTTAATCCCTTGTCTGTC	2146
Db	1496	GAATACACAGGATATTTTGGGGCGAGAGATATAAACGTTAGTTAATCCCTTGTCTGTC	1555
QY	2147	AATCAACAGTCTCAGTTCTCTTGTCTTTCAATGTACTTAAACCTCCTGCTGTGGCTCGCA	2206



```
Db 1556 AATCAGAGTCTCAGTCTCTTCTGCTTTTCCACATTTGTAATTAACCTTCCCTGCTGCTCGCA 1615
Qy 2207 TCCTATCTTAATAAAGAACATCTTGAATATCAAAAAAAGAAAAA 2260
Db 1616 TCCTAGCTTAATAAAGAACATCTTGAATATCAAAAAAAGAAAAA 1669

RESULT 12
AF395124 1625 bp mRNA linear PRI 28-OCT-2003
LOCUS Homo sapiens WWOXdelta5-8 mRNA, complete cds, alternatively
DEFINITION spliced.
ACCESSION AF395124
VERSION AF395124.1 GI:15028464
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1625)
AUTHORS Bednarek,A.K., Keck-Waggoner,C.L., Daniel,R.L., Laflin,K.J.,
Bergsagel,P.L., Kiguchi,K., Brenner,A.J. and Aldaz,C.M.
TITLE WWOX, the FRA16D gene, behaves as a suppressor of tumor growth
JOURNAL Cancer Res. 61 (22), 8068-8073 (2001)
MEDLINE 21575822
PUBMED 11719429
REFERENCE 2 (bases 1 to 1625)
AUTHORS Bednarek,A.K. and Aldaz,C.M.
TITLE Direct Submission
JOURNAL Submitted (25-JUN-2001) Carcinogenesis, The University of Texas,
M.D. Anderson Cancer Center, Science Park-Research Division, Park
Road 1C, Smithville, TX 78957, USA
LOCATION/Qualifiers
FEATURES
source 1..1625
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
134..1069
/notes="aberrant WWOX transcript found in cancer cells"
CDS
/evidence=not_experimental
/product="WWOXdelta5-8"
/protein_id="AAK81728.1"
/db_xref="GI:15028465"
/translations="MAALRYAGLDDTDSDELPPGWEERTTKDGVVYANHTEKQW
EHPKTKRKEVAGDLPGHEQEDENGQVFFVDHINKRTTYLDPLAFTVDNPTKPT
TRQYDGTAMAILGRDFIGKVVVYTGANSIGATGSCHHRVLCCTPRFGSGRDVL
QQLPLHALTRSSRDGPDVGAQREADPTAWQVRLSGAQSQAQWHAFTALCVSPA
SARAGPLPNVPTQIIRKSKNSHNRKVLKLYQWEAGNSWGKVSFLFWGWARHSLCF
LVVACLKVKTCLVCRFRISLEKHQQSFYCYRIA"

ORIGIN
Query Match 47.9%; Score 1084; DB 9; Length 1625;
Best Local Similarity 100.0%; Pred. No. 1e-285;
Matches 1084; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1181 GCACAGGAGCTGCCACACCGGTACTGTGCTGCTGCCAGAACTCAGAGGCTCTGGG 1240
Db 542 GCACAGGAGCTGCCACACCGGTACTGTGCTGCTGCCAGAACTCAGAGGCTCTGGG 601

Qy 1241 AGGAGTACTTCAACAACTGCTGCCGCTGCATGCCCTCACCAAGCTCAGAGCGAAGA 1300
Db 602 AGGAGTACTTCAACAACTGCTGCCGCTGCATGCCCTCACCAAGCTCAGAGCGAAGA 661

Qy 1301 GACGCCCGACCCCTGTGGGGCTCAGCAGAGGCTGATCAAGAACGGCTTGGCAGCCA 1360
Db 662 GACGCCCGACCCCTGTGGGGCTCAGCAGAGGCTGATCAAGAACGGCTTGGCAGCCA 721

Qy 1361 GTCGGGTAGTGAGCTCAGAGCGGATGGCACAACACCCGCTGTGTGTCCTT 1420
Db 722 GTCGGGTAGTGAGCTCAGAGCGGATGGCACAACACCCGCTGTGTGTCCTT 781
```

```
Qy 1421 CACGCAAGTCCCAAGGCTGGGCCCTTCCAAATGTTCCTCCAAACACAGATCCGCAAGAGT 1480
Db 782 CACGCAAGTCCCAAGGCTGGGCCCTTCCAAATGTTCCTCCAAACACAGATCCGCAAGAGT 841

Qy 1481 AAAGGAATTAAGAGCAGTCAACACAGAGTGAATAATCTTAAGTACCATGGGAAGCAGGG 1540
Db 842 AAAGGAATTAAGAGCAGTCAACACAGAGTGAATAATCTTAAGTACCATGGGAAGCAGGG 901

Qy 1541 AATTCTGGGGTAAAGTATCACTTTCTGGGGCTGGGCTAGGCATAGTCTCTTTGCTTT 1600
Db 902 AATTCTGGGGTAAAGTATCACTTTCTGGGGCTGGGCTAGGCATAGTCTCTTTGCTTT 961

Qy 1601 CTGGTGGTCCCTGTTTGAAGTAAAGAACTCTGCTGTGTGTGTAGTTCCTGATCTCCCTG 1660
Db 962 CTGGTGGTCCCTGTTTGAAGTAAAGAACTCTGCTGTGTGTGTAGTTCCTGATCTCCCTG 1021

Qy 1661 GAGAAGCACACAGCAATCTCTTTCTTTTATAGTATAGCTCAGCTCCCTCGTC 1720
Db 1022 GAGAAGCACACAGCAATCTCTTTCTTTTATAGTATAGCTCAGCTCCCTCGTC 1081

Qy 1721 CCATCCAGCTTACCACACGCGCCACCTGTCAGCGCGGGGCTGGCTTCTCTACTTAGGG 1780
Db 1082 CCATCCAGCTTACCACACGCGCCACCTGTCAGCGCGGGGCTGGCTTCTCTACTTAGGG 1141

Qy 1781 AAAAAAAGCAAGTGTCTACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1840
Db 1142 AAAAAAAGCAAGTGTCTACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1201

Qy 1841 CATCTGACCAAGCTGAGCCAGCTTAGCACTGCTGGGGGAGACAAATCTCAGAACCTTG 1900
Db 1202 CATCTGACCAAGCTGAGCCAGCTTAGCACTGCTGGGGGAGACAAATCTCAGAACCTTG 1261

Qy 1901 TCCAGCCAGTGAAGATGACAGTGACACCCAGAGGAGTAGAATACCGAAGCTACCCAGG 1960
Db 1262 TCCAGCCAGTGAAGATGACAGTGACACCCAGAGGAGTAGAATACCGAAGCTACCCAGG 1321

Qy 1961 TGGCAAAAGTACTGTGTATAGACTCTTTGCTTAATGTATGCTGCTGCTGCTGCTGCTGCT 2020
Db 1322 TGGCAAAAGTACTGTGTATAGACTCTTTGCTTAATGTATGCTGCTGCTGCTGCTGCTGCT 1381

Qy 2021 ATAACAAATTTTCAATCAATCTTCTAGATACCTTGAAGCGAGGAGGAGGAGGCTATAT 2080
Db 1382 ATAACAAATTTTCAATCAATCTTCTAGATACCTTGAAGCGAGGAGGAGGAGGCTATAT 1441

Qy 2081 ACTTAAGAAATACACAGGATATTTGGGGGCGAGAGATAAAGCTTAGTTAATCCCTTTG 2140
Db 1442 ACTTAAGAAATACACAGGATATTTGGGGGCGAGAGATAAAGCTTAGTTAATCCCTTTG 1501

Qy 2141 TCTGTCAATCACAGTCTCAGTCTCTTTGCTTTTCAATGTATTAACCTCTGCTGCTGCTG 2200
Db 1502 TCTGTCAATCACAGTCTCAGTCTCTTTGCTTTTCAATGTATTAACCTCTGCTGCTGCTG 1561

Qy 2201 CTCGCATCTCTATCTTAATAAAGAACATCTTGAATATCAAAAAAAGAAAAA 2260
Db 1562 CTCGCATCTCTATCTTAATAAAGAACATCTTGAATATCAAAAAAAGAAAAA 1621

Qy 2261 AAAA 2264
Db 1622 AAAA 1625

RESULT 13
AF187014
LOCUS Mus musculus WW-domain oxidoreductase mRNA, complete cds.
DEFINITION
ACCESSION AF187014
VERSION AF187014.1 GI:6934273
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 2197)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
```

AUTHORS Chang, N. S., Pratt, N., Heath, J., Schultz, L., Sleeve, D., Carey, G. B.  
and Zevotek, N.  
TITLE Hyaluronidase induction of a WW domain-containing oxidoreductase  
that enhances tumor necrosis factor cytotoxicity  
J. Biol. Chem. 276 (5), 3361-3370 (2001)  
JOURNAL MEDLINE 21264809  
PUBMED 11058590  
REFERENCE 2 (bases 1 to 2197)  
AUTHORS Chang, N.-S.  
TITLE Direct Submission  
JOURNAL Submitted (15-SEP-1999) Laboratory of Molecular Immunology, Guthrie  
Research Institute, 1 Guthrie Square, Sayre, PA 18840, USA  
FEATURES  
source  
1. .2197  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
72. .1316  
/codon\_start=1  
/product="WW-domain oxidoreductase"  
/protein\_id="AAF31693.1"  
/db\_xref="GI:6934274"  
/translation="MAALRYAGLDITDSEDELPPCHERTTYDGVNYYVANHTEETQW  
EHPKTKRKRVAGDLPYGEQEDENGQVFDHINKRTTYLDPLAFVDDNPDKPT  
TRQYDGSSTAMEILQGRDFTGKVLVTGANSIGIFETAKSPALGHVILACRLSR  
ASEAVSRILEEWHKAKVAMTLDLAVRSVQHFPAFAKAKNVLSHLVDCINAGTFLPW  
GLTKDGLLETFQANHLGHFVLQDLVLCRSSPARVIVSSESHRFDINDSSKGLD  
LSRLSPRSDYVAMLYNRSKLNILFSLNELHRLSPRGVSNVHPGNMYSATHRN  
SWYKLLPLARPPFKSQOQKATTVYCAVAPELGLGMVFNRCRLPSEASAE  
TARALWEUSERLIQRLGSPSS"  
ORIGIN  
Query Match 47.6%; Score 1077; DB 10; Length 2197;  
Best Local Similarity 73.7%; Pred. No. 9.2e-284;  
Matches 1524; Conservative 0; Mismatches 490; Indels 53; Gaps 10;  
53 GGCTACTTCTGCTCCAGCATGGCAGCTCTGGCTATGCGGGCTGGAGCAGCAGGACAG 112  
107 GTGCTCTCACAGTACAGCATGGCAGCTGGCTGCGCTACGCGGGCTGGAGCAGCAGGACAG 166  
167 TGAGGAGAGCTGCTCCGGCTGGAGAGAGAACACCAAGAGCGCTGGTTACTA 226  
113 TGAGGATGAGCTGCTCCGGCTGGAGAGAGAACCAAGAGCGCTGGTTACTA 172  
227 CGCCATCACCGAGAGAGAGACTCAGTGGGAAACATCCAAATCTGGAAGAAAGAG 286  
173 CGCCATCACCTGAGAGAGAGACCACTGGGAAACATCCGAACCGGAGAGGAGAG 232  
287 AGTGGCAGGAGATTGCGCATAGGATGGGAAACAGAGAACTGATGAGAACCGAGAGTGT 346  
233 GGTGCGCAGGAGATTGCGCATAGGATGGGAAACAGAGAGACTGAGAGAACCGAGAGTGT 292  
347 TTTTGTGACCATATAAATAAGACCACTTCTTGGACCAAGACTGGCGTTTACTGT 406  
293 TTTGTTGACCATATAAATAAGACCACTTCTTGGATCCAGATTGGCATTTACCGT 352  
407 GGATGATAATCGGACCAAGCAACCAACCGGAAAGATACGAGCGAGCACCACCTGCCAT 466  
353 GGACGATATCCGACGAGCCGACCAACAGAGATACGAGCGAGCACCACCGCAT 412  
467 GGAATTCCTCAGGCGCGGATTCATCGCAAAGTGGTTGGTCACTGGAGCTAATTC 526  
413 GGAGATCTCTCAGGCGCGGACTTCAAGCAAGTGGTGGTCTGTTACTGGAGCAATTC 472  
527 AGGAATAGGGTTCGAACCGCAAGTCTTTTGCCTCCATGGTGGCATGTGATCTTGGC 586  
473 GGAATAGGGTTCGAGACTGCCAAGTCTTTTGCCTCCATGGCGCGCATGTGATCTTGGC 532  
587 CTGAGGAACATGGCAAGCGCGAGTGAAGCAGTGTCAACGATTTTGAAGAAGTGGCATAA 646  
533 CTGTAGGAACCTGAGCAGAGCAGCGAGCGCGTGTCCGATCTCTGGAGAGTGGACAA 592  
647 AGCCAAGGTAGAGCAATGACCCCTGGACCTCGCTCTGCTCCGTPAGCGTGCAGCATTTTGC 706

593 AGCCAAGGTAGAGCAATGACCCCTGGACCTGACCTGCTCCGCTAGCGTGCAGCATTTTGC 652  
707 TGAACATTTCAAGGCCAAGAATGTCCTCTTCATGTGCTTGTGTGCAAGCAGCAACTTT 766  
653 CGAAGCATTTCAAGGCCAAGAATGTCCTCTTCACGTGCTTGTGTGTAATGCAAGGACGTT 712  
767 TGCTCTACCTGGAGCTCTCACCAAGATGGCTGGAGACCACTTTCAAGTGAATCATCT 826  
713 TGCTCTGCTGGCGCTTCAAGAGATGGCTGGAGCAACCTTTCAAGCGCAACCATCT 772  
827 GGGGCACTTTACCTTGTCCAGCTCTCTCAGGATGTTTTGTGCGCTCAGCTCTCTGCCCG 886  
773 GGGACATTTTACCTTGTCCAACTCTCGAGATGTTCTGTGCGCTCATCCCCCGCG 832  
887 TGTCAATTGTGCTCTCTCAGAGTCCCATGATTTACAGATATTAAGCATCTCTTGGGAAA 946  
833 TGTCAATCGTGTGCTCTTCCGAGTCCCAAGATTTACAGATATTAATGATCTCTCCGGHAA 892  
947 ACTGGACTTCAGTGGCTCTCTCCAAACAAACAAACAGCTATTGGGGGATGCTGGCTTATAA 1006  
893 ACTTGACCTCAGCGCGCTGTCTCCACGAGGAGCGACTACTGGGCCATGCTGGCTTATAA 952  
1007 CAGGTCCAAGCTCTCCAAATCTCTTCTCCAAAGCTGCAACCGTTCGCTCTCCCCAG 1066  
953 CAGATCTAACTCTGCAATATCTCTTCTCCAAATGAACTGCACCGCTGGCTCTCCCCCG 1012  
1067 CGGGCTCAGCTCGAAGCAGTGCATCTTGGAAATATGATGATCTCAACATCATTCGAG 1126  
1013 CGGGGTCACTCCAAAGCAGTGCATCTTGGGAAATGATGATGATGATGATGATGATGATG 1072  
1127 CTGGTGGGTGTACACACTGCTGTTTACCTTGGCGAGGCTTTTCAACAAAGTCCATGCAAC 1186  
1073 CTCATGGGTGTACAGCTGCTGTTCACTTGGCAGCGCTTCAACAAAGTCCATGCAAC 1132  
1187 GGGAGTGCACCAACCGTGTACTGTGCTGTCTCCAGAACTTGGAGGGTCTGGAGGGAT 1246  
1133 GGGAGTGTACCACTGTCTACTGTGCTGCTGCCCGCAGAGCTGGAGGGCTCGAGGGAT 1192  
1247 GTACTTCAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1306  
1193 GTATTTCACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1252  
1307 CGGAGCCCTGTGGCGCTCAGCAGAGGCTGATCAAGAAACGGCTTGGCAGCGAGTCCGG 1366  
1253 TCGGGCCCTGTGGAACTCAGTGAGGCTCATCCAGGACCGACTGGGTAGTCCATCCAG 1312  
1367 CTAAATGGAGCTCAGAGCGGATGGGACACACACCGCGCTGTGTGTGCTGCTGCTGCTGCT 1426  
1313 CTAAATGGAAACCCATAATATTGGGCAAAACCCACCCATGCTGCTGCTGCTGCTGCTGCT 1372  
1427 AGTGCAGGGCTGGCGCTTCCAAATGCTCCCAACACAGATCCGCAAGAGTAAAGGA 1486  
1373 CAGCCC-----CATCATCTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1423  
1487 AATAAGAGCTCACAACAGAGTGAATAATTTAAGTACCAATGGGAAGCAGAGGAAATTC 1546  
1424 GATAAGAACAAATCAA-----TGGGAGAGGTATAACTAGTGGGAGCAGGAAATTC 1476  
1547 TGGGTAAAGTATCATTTTCTGGGGCTGGGCTAGGCATAGTCTCTTTGC-----TTTCTGCT 1605  
1477 AAGAGATGA-TGTCCCTTTCTAGGACTGGGTAGTCTGGGAGCTGCTTGTCTTTCTGCT 1535  
1606 GGTGCTCTTTGAAGTAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1665  
1536 GTTGGCCGCTTGAATAAGAAATGGAATCATGTTAGTGAATGATGATGATGATGATGATG 1595  
1666 GCACGAGCAATTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1725  
1596 GCACAGCTGATTTCTAT-----CAACCTTAGAATAACCTGATATCCCTCGACCCACC 1649  
1726 CAGCTACCAACCGGCGACCACTGAGCGGGGGCTGGCTTCTCTCTTCTTCTTCTTCTTCTT 1785

Db 1650 CAGCTGCCACCACTGCCAATTTGCCACCACCAAGGTGGCTTTCTCTCTTAATGGAAGAA 1709  
 QY 1786 AAAGCA-----AGTGTTCACCTGCTCTCTGCTGATTCAGGAGATTAATGTT 1835  
 Db 1710 AAAAAAACACACACATTTCTCTCTTCCAGCTGAGTTCAATCTCTGGAGCCATGTT 1769  
 QY 1836 TCATTCATCTGACCAAGATGACCCAGCTTTAGCAATCTGCTGGGAGACAAATCTCGAA 1895  
 Db 1770 TCACATTCCTTACCAAGTCTGACGAGCGTGGCACTTGAGAG-----AAAACTCAGTA 1825  
 QY 1896 CTTGTGTC-----AGCCAGTGAAGATGACAGTGCACCCAGGAGGAGTAGAATA 1945  
 Db 1826 CTTAGTCCCCAAGAGAACCAATGCGACGCTGCCATCAGAGAGGACGATCGGTGGTG 1885  
 QY 1946 CGCAGAACTACCAAGTGGC--AAAGTACTTGTCTATAGACTCTTTGCTAATGCTATGCAA 2003  
 Db 1886 TGCAGACTGGTGTTCAGAGACACATTCCTGACCTTCCTGCCAAGTTAAACAA 1945  
 QY 2004 AAAATCTTTAGAGATTATAACAAAT---TTTTCAATCATTCCTTAGATACCTTGAAG 2060  
 Db 1946 GCAATCTTTAGAGATTATAACGAGTGCATTTTACATTTTGCAAAACCATTCCTTAGGCAG 2005  
 QY 2061 GCAGGAGGAGGAGGTATATACITTAAGATACACAGGATTTTGGGGGCGACAGATAA 2120  
 Db 2006 GCAAGGAGGAGGACGCTGCTCTTAGATATACATGGGATATGCTGGAGGTGGGAAGCAAG 2065  
 QY 2121 AACGTTAGTTAAATCCCTTTGCTGTGCA 2147  
 Db 2066 GAAGTAGTCTCCCTCCAGTCTTCACTCA 2092

RESULT 14  
 AC092376 168083 bp DNA linear PRI 15-MAR-2003  
 LOCUS Homo sapiens chromosome 16 clone RP11-679B19, complete sequence.  
 DEFINITION AC092376 AC036142  
 ACCESSION AC092376.3 GI:28973808  
 VERSION HTG.  
 KEYWORDS Homo sapiens (human)  
 SOURCE  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 168083)  
 AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los Alamos National Laboratory.  
 TITLE Direct Submission  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 168083)  
 AUTHORS DOE Joint Genome Institute.  
 JOURNAL Direct Submission  
 REFERENCE 3 (bases 1 to 168083)  
 AUTHORS DOE Joint Genome Institute.  
 JOURNAL Submitted (03-JUL-2001) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 REFERENCE 4 (bases 1 to 168083)  
 AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los Alamos National Laboratory.  
 JOURNAL Submitted (15-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 COMMENT On Mar 15, 2003 this sequence version replaced gi:15187320.  
 Draft Sequence Produced by DOE Joint Genome Institute  
 www.jgi.doe.gov  
 Finishing Completed at Stanford Human Genome Center and Los Alamos National Laboratory  
 www.shgc.stanford.edu  
 Quality: Phrap Quality >=40 100% of Sequence;  
 Estimated total Number of Errors is 0.  
 Location/Qualifiers  
 1..168083

/organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="16"  
 /clone="RP11-679B19"

ORIGIN

Query Match 46.9% Score 1062.2; DB 9; Length 168083;  
 Best Local Similarity 99.3%; Pred. No. 2.8e-279;  
 Matches 1067; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1167 TTCAACAAGTCCATGCAACAGGAGCTGCCACCAACCGTGTACTGTGTGCTGCCAGAA 1226  
 Db 23701 TTCTTGATTTTCAGCAACAGGAGCTGCCACCAACCGTGTACTGTGTGCTGCCAGAA 23760  
 QY 1227 CTGAGAGGTCTGGAGGAGTGTACTTCAACAACTGTGCGCTGCATGCCCTCAACGAA 1286  
 Db 23761 CTGAGAGGTCTGGAGGAGTGTACTTCAACAACTGTGCGCTGCATGCCCTCAACGAA 23820  
 QY 1287 GCTCAGAGCAAGAGAGCGCCGACCTGTGGCGCTCAGCGAGAGGCTGATCCAGAA 1346  
 Db 23821 GCTCAGAGCAAGAGAGCGCCGACCTGTGGCGCTCAGCGAGAGGCTGATCCAGAA 23880  
 QY 1347 CGCTTTGGCAGCAGTCCGGCTAAGTGGAGCTCAGAGCGGATGGGCAACACACCCGCC 1406  
 Db 23881 CGCTTTGGCAGCAGTCCGGCTAAGTGGAGCTCAGAGCGGATGGGCAACACACCCGCC 23940  
 QY 1407 TGTGTGTCTCCCTCAGCAAGTGCAGGCTGGGCGCTTCCAAATGTCCCTCCACAC 1466  
 Db 23941 TGTGTGTCTCCCTCAGCAAGTGCAGGCTGGGCGCTTCCAAATGTCCCTCCACAC 24000  
 QY 1467 AGATCCGCAAGAGTAAAGNAATAAGAGAGTCAACACAGAGTGAATAATCTTAAGTACC 1526  
 Db 24001 AGATCCGCAAGAGTAAAGNAATAAGAGAGTCAACACAGAGTGAATAATCTTAAGTACC 24060  
 QY 1527 AATGGAAGCAGGGAATCTCTGGGGTAAAGTATCATCTTTCTGGGCTGGGCTAGGCATA 1586  
 Db 24061 AATGGAAGCAGGGAATCTCTGGGGTAAAGTATCATCTTTCTGGGCTGGGCTAGGCATA 24120  
 QY 1587 GGTCTCTTTCTTTCTGGTGGCTGTGTTGAAAGTAAAAACCTCTTGGTGTAGGT 1646  
 Db 24121 GGTCTCTTTCTTTCTGGTGGCTGTGTTGAAAGTAAAAACCTCTTGGTGTAGGT 24180  
 QY 1647 TCGCTATCTCCCTGGAGAGCACCAGCAATCTCTTTCTTTTACTGTATAGAAATAGCCT 1706  
 Db 24181 TCGCTATCTCCCTGGAGAGCACCAGCAATCTCTTTCTTTTACTGTATAGAAATAGCCT 24240  
 QY 1707 GAGTCCCTCTCGTCCCATCCAGTACCAACACGCGCCACCACTGCAGCGGGGCTGGCCT 1766  
 Db 24241 GAGTCCCTCTCGTCCCATCCAGTACCAACACGCGCCACCACTGCAGCGGGGCTGGCCT 24300  
 QY 1767 TCTCTACTTAGGAAGAAAAAGCAAGTGTTCATCTCTCTCTGCTGCATTTGATCCAGGAG 1826  
 Db 24301 TCTCTACTTAGGAAGAAAAAGCAAGTGTTCATCTCTCTCTGCTGCATTTGATCCAGGAG 24360  
 QY 1827 ATAAATGTTTCAATCTGACCAAGACTGAGCCAGCTTAGCAATCTGCTGGGAGACAA 1886  
 Db 24361 ATAAATGTTTCAATCTGACCAAGACTGAGCCAGCTTAGCAATCTGCTGGGAGACAA 24420  
 QY 1887 ATCTCAGAACTTGTCTCCAGCCAGTGAAGTACAGTGCACCCAGAGGAGTAGAATAC 1946  
 Db 24421 ATCTCAGAACTTGTCTCCAGCCAGTGAAGTACAGTGCACCCAGAGGAGTAGAATAC 24480  
 QY 1947 GAGAACTACCAAGTGGCAAGTACTTGTTCATAGACTCTTTGCTTAATCTATGCAAAA 2006  
 Db 24481 GAGAACTACCAAGTGGCAAGTACTTGTTCATAGACTCTTTGCTTAATCTATGCAAAA 24540  
 QY 2007 ATTCTTTAGAGATTAAACAAATTTTCAAACTTCCTCTAGATACCTTTGAAGGAGGA 2066  
 Db 24541 ATTCTTTAGAGATTAAACAAATTTTCAAACTTCCTCTAGATACCTTTGAAGGAGGA 24600  
 QY 2067 AGGGAAGCGTATATATCTTAAGAAATACACAGGATATTTTGGGGGCGACAGAAATAACGTT 2126



```
QY 1587 GGTCCTTTGCTTTCTGGTGGGCTGTTTCAAGTAAAAACCTGCTTGTGTGTAGGT 1646
Db |||||
QY 69052 GGTCCTTTGCTTTCTGGTGGGCTGTTTCAAGTAAAAACCTGCTTGTGTGTAGGT 68993
Db |||||
QY 1647 TCCGTATCTCCCTGAGAGACACGACGATTCCTCTTTTCTTTTACTGTTATAGATAGCCT 1706
Db |||||
QY 68992 TCCGTATCTCCCTGAGAGACACGACGATTCCTCTTTTCTTTTACTGTTATAGATAGCCT 68933
Db |||||
QY 1707 GAGGTCCTCTGCTCCCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 1766
Db |||||
QY 68932 GAGGTCCTCTGCTCCCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 68873
Db |||||
QY 1767 TCTCCTACTTAGGGAAGAAAAAGCAAGTGTTCATCTGCTCCTTTGCTGCAATGATCCAGGAG 1826
Db |||||
QY 68872 TCTCCTACTTAGGGAAGAAAAAGCAAGTGTTCATCTGCTCCTTTGCTGCAATGATCCAGGAG 68813
Db |||||
QY 1827 ATAATTTGTTTCAATTCATCTGACCCAGAGCTGAGCCAGCTTAGCACTGCTGGGGAGACAA 1886
Db |||||
QY 68812 ATAATTTGTTTCAATTCATCTGACCCAGAGCTGAGCCAGCTTAGCACTGCTGGGGAGACAA 68753
Db |||||
QY 1887 ATCTCAGAACCTTTGCTCCAGCCAGTGCAGATGACAGTGCACACCCAGAGGGAGTAGAATAC 1946
Db |||||
QY 68752 ATCTCAGAACCTTTGCTCCAGCCAGTGCAGATGACAGTGCACACCCAGAGGGAGTAGAATAC 68693
Db |||||
QY 1947 GCAGAACTACGAGTGGCAAGTACTTGTCTATAGACTCCTTTGCTTAATGCTATGCAAAA 2006
Db |||||
QY 68692 GCAGAACTACGAGTGGCAAGTACTTGTCTATAGACTCCTTTGCTTAATGCTATGCAAAA 68633
Db |||||
QY 2007 ATCTTTAGAGATTATACAAATTTTCAATTCATCTGATACCTTTGATACCTTTGAAGGCAGGA 2066
Db |||||
QY 68632 ATCTTTAGAGATTATACAAATTTTCAATTCATCTGATACCTTTGATACCTTTGAAGGCAGGA 68573
Db |||||
QY 2067 AGGGAAGCGTATATACCTTAAGAAATACAGAGATATTTTGGGGGGCAGAGATATAAACGTT 2126
Db |||||
QY 68572 AGGGAAGCGTATATACCTTAAGAAATACAGAGATATTTTGGGGGGCAGAGATATAAACGTT 68513
Db |||||
QY 2127 AGTTAATCCCTTTGCTGCTCAATCAAGTCTCAGTTCTCTTGTCTTCAATTTGTTACTTTAA 2186
Db |||||
QY 68512 AGTTAATCCCTTTGCTGCTCAATCAAGTCTCAGTTCTCTTGTCTTCAATTTGTTACTTTAA 68453
Db |||||
QY 2187 ACCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2241
Db |||||
QY 68452 ACCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 68398
Db |||||
```

Search completed: May 25, 2004, 03:10:36  
Job time : 8869 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 09:49:46 ; Search time 58 Seconds  
(without alignments)  
2016.805 Million cell updates/sec

Title: US-09-978-318B-1  
Perfect score: 2203  
Sequence: 1 MAALRYAGLDDTDEDELP.....RTLWALSERLIQERLGSQSG 414

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues  
Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*  
1: Geneseq1980s.\*  
2: Geneseq1990s.\*  
3: Geneseq2000s.\*  
4: Geneseq2001s.\*  
5: Geneseq2002s.\*  
6: Geneseq2003as.\*  
7: Geneseq2003bs.\*  
8: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2203	100.0	414	5	ABG96559 Human sho
2	2196	99.7	414	4	AAB93119 Human pro
3	2196	99.7	414	4	AAB85886 Human gen
4	2191	99.5	414	4	AAB84485 Amino aci
5	2190	99.4	415	4	AAB18308 Human end
6	2184	99.1	423	4	ABB10203 Human CDN
7	2184	99.1	423	5	ABP66790 Human pol
8	1869	84.8	363	4	AAB84484 Amino aci
9	1023	46.4	409	4	ABG63100 Drosophil
10	927	42.1	189	4	AAB84486 Amino aci
11	862.5	39.2	173	4	ABBI10434 Human CDN
12	862.5	39.2	173	4	AAB18436 Human end
13	862.5	39.2	173	5	ABP67021 Human pol
14	550.5	25.0	316	5	AAE14438 Human dru
15	549.5	24.9	316	6	ADA54192 Human pro
16	549.5	24.9	316	7	ABR82985 Human unu
17	540	24.5	321	5	ABP69268 Human pol
18	533.5	24.2	322	3	AAG48405 Arabidops
19	527	23.9	323	2	AAW15759 Cotton fi
20	520.5	23.6	320	3	AAQ05338 Arabidops
21	519.5	23.6	296	4	ABB58831 Drosophil
22	514.5	23.4	316	3	AAG48393 Arabidops
23	513.5	23.3	318	3	AAy82014 Human imm
24	513.5	23.3	318	3	AAB12158 Hydrophob
25	513.5	23.3	318	4	AAm01129 Human pro

## ALIGNMENTS

## RESULT 1

ABG96559 standard; protein; 414 AA.

AC ABG96559;

DT 12-DEC-2002 (first entry)

DE Human short chain dehydrogenase family member WWOX.

Short chain dehydrogenase; SDR; human; antimitocotic; pesticide; herbicide; DHPR deficiency; phenylketonuria; galactosaemia III; diethyl CoA reductase deficiency; adrenal hyperplasia; ovarian cancer; adrenogenital syndrome; mineralcorticoid excess syndrome; breast cancer; male pseudohemaphroditism; Zellweger syndrome; Down's syndrome; polycystic kidney disease; Alzheimer's disease; retinitis pigmentosa; retinitis punctata albescens; arterial hypertension; follicular lymphoma; hepatocarcinogenesis; fungicide; antibiotic.

OS Homo sapiens.

PN WO200212544-A2.

PD 14-FEB-2002.

PF 07-AUG-2001; 2001WO-EP009140.

PR 07-AUG-2000; 2000US-0223436P.

PA (BION-) BIONETWORKS GMBH.

PI Wilckens T;

DR WPI; 2002-241770/29.

Identifying or verifying members of the short chain dehydrogenase (SDR) family, useful for novel drug development (e.g. for the development of antimitocotic, pesticides or herbicides), by employing an algorithm using core SDR motifs.

Disclosure; Fig 4; 168pp; English.

The invention relates to identifying or verifying members of the short chain dehydrogenase (SDR) family comprising employing an algorithm using core SDR motifs (MFI-MR4 and MVI, WV2 given in the specification) for searching members of the SDR family. Also included are a member of the SDR family identified with the method above, a method for providing modulators for members of the SDR family, a method for evaluation of lead

Aam93777 Human pol  
Aau69774 Human pro  
Aab74812 Prostate  
Aag99014 Human pro  
Abu71665 Prostate  
Abg96558 Human sho  
Abb95234 Human P50  
Adb54346 Prostate  
Adb75403 Prostate  
Adb13789 Human pro  
Aay48616 Human bre  
Aab36900 Human ARS  
Abr63833 Human sho  
Abr82986 Human CGI  
Abb58826 Drosophil  
Aag81281 Human AFP  
Abb66596 Drosophil  
Aay41730 Human PRO  
Aab44286 Human PRO  
Aau29114 Human PRO

26 513.5 23.3 318 4 AAM93777  
27 513.5 23.3 318 4 AAU69774  
28 513.5 23.3 318 4 AAB74812  
29 513.5 23.3 318 4 AAG99014  
30 513.5 23.3 318 4 ABU71665  
31 513.5 23.3 318 5 ABG96558  
32 513.5 23.3 318 5 ABB95234  
33 513.5 23.3 318 6 ADB54346  
34 513.5 23.3 318 7 ADB75403  
35 513.5 23.3 318 7 ADB13789  
36 513.5 23.3 333 2 AAY48616  
37 512.5 23.3 314 3 AAB36900  
38 512.5 23.3 318 7 ABR63833  
39 512.5 23.3 318 7 ABR82986  
40 510.5 23.2 300 4 ABB58826  
41 508 23.1 296 4 AAG81281  
42 507 23.0 594 4 ABB66596  
43 505 22.9 336 2 AAY41730  
44 505 22.9 336 3 AAB44286  
45 505 22.9 336 4 AAU29114







CC associated with the tumour, a breakpoint of one or more chromosomal  
CC rearrangements associated with the tumour and a pause site within the  
CC FRA16 gene. FOR nucleic acid molecules are useful as markers to identify  
CC relationship between the fragile site (FRA16D) and the DNA instability in  
CC neoplasia which allows better diagnosis of cancers associated with the  
CC region  
XX  
SQ Sequence 414 AA;

Query Match 99.5%; Score 2191; DB 4; Length 414;  
Best Local Similarity 99.5%; Pred. No. 1.7e-214;  
Matches 412; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAALRYAGLDDTSDSELPDPCWEERTTKDGMVYVYANHTTEKTQWEHPKTKGRKRVAGDLP 60  
Db 1 MAALRYAGLDDTSDSELPDPCWEERTTKDGMVYVYANHTTEKTQWEHPKTKGRKRVAGDLP 60

Qy 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDTKPTTRQRYDGTSTAMEILQGR 120  
Db 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDTKPTTRQRYDGTSTAMEILQGR 120

Qy 121 DFTGKVVVVTGANSIGIGETAKSPALHGAHVILACRNWARASEAVSRILEEWHKAKVEM 180  
Db 121 DFTGKVVVVTGANSIGIGETAKSPALHGAHVILACRNWARASEAVSRILEEWHKAKVEM 180

Qy 181 TIDLALLRSVQHFAFAKAKNVPLHLVCNATFALPWSLTKDGLTTFQVNHHLGHFYL 240  
Db 181 TIDLALLRSVQHFAFAKAKNVPLHLVCNATFALPWSLTKDGLTTFQVNHHLGHFYL 240

Qy 241 QLLQDVLCSAPARIVIVSSSHRFTDINDSLGKLDLFRSLTKNDYWLAMLYNRSKLCN 300  
Db 241 QLLQDVLCSAPARIVIVSSSHRFTDINDSLGKLDLFRSLTKNDYWLAMLYNRSKLCN 300

Qy 301 ILFSNELHRLSPRGVTSNAVHPGNMYSNHRSMWVYTLFTLARPPTKSMQOQAATTV 360  
Db 301 ILFSNELHRLSPRGVTSNAVHPGNMYSNHRSMWVYTLFTLARPPTKSMQOQAATTV 360

Qy 361 YCAAPELEGLGGMVFNCCRCMPSPQAQSEBTARTLWALSERLQERLGSQSG 414  
Db 361 YCAAPELEGLGGMVFNCCRCMPSPQAQSEBTARTLWALSERLQERLGSQSG 414

RESULT 5  
AAU18308  
ID AAU18308 standard; protein; 415 AA.  
AC AAU18308;  
XX 21-NOV-2001 (first entry)  
DE Human endocrine polypeptide SEQ ID No 263.  
XX Endocrine protein; human; mouse; rabbit; goat; horse; food additive; cat;  
KW dog; chicken; sheep; immunosuppressive; antiarthritic; vasotrophic;  
KW antirheumatic; antiproliferative; cytostatic; cardiac; neuroprotective;  
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;  
KW ophthalmological; vulnary; gene therapy; autoimmune disease; neoplasm;  
KW hyperproliferative disorder; breast; liver; cardiovascular disorder;  
KW cerebrovascular disorder; nervous system disorder; bacterial infection;  
KW fungal infection; viral infection; ocular disorder; endocrine disorder;  
KW gastrointestinal disorder; renal disorder; respiratory disorder;  
KW wound healing; skin aging; organ transplantation; food preservative;  
KW tissue regeneration; anti-infertility.  
XX Homo sapiens.  
XX WO20015364-A2.  
PN  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US001308.  
PF  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR



XX  
OS Homo sapiens.  
XX WO200154474-A2.  
PN  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001349.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0218963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 23-AUG-2000; 2000US-0227182P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
XX  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 11-DEC-2000; 2000US-0251990P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX



XX The invention relates to novel genes (ABV83682-ABV84101) and proteins  
CC (ABP66710-ABP67129) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing  
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
CC infectious diseases such as viral, bacterial, fungal and parasitic  
CC infections. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 423 AA;

Query Match 99.1%; Score 2184; DB 5; Length 423;  
Best Local Similarity 99.0%; Pred. No. 9.2e-214; Indels 0; Gaps 0;  
Matches 410; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MAALRYAGLDGTTSEDELPDGGWEERTTKDGVVYANHTTEKTQWHPKTKGKRKRVAGDLP 60  
DB 10 MAALRYAGLDGTTSEDELPDGGWEERTTKDGVVYANHTTEKTQWHPKTKGKRKRVAGDLP 69  
QY 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDNPTKPTRQRYDGSITAMAILQGR 120  
DB 70 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDNPTKPTRQRYDGSITAMAILQGR 129  
QY 121 DFTGKVVVVTGANGSIGFETAKSFALHGAHVILACNNMARASEAVSRILEEWHKAKVETM 180  
DB 130 DFTGKVVVVTGANGSIGFETAKSFALHGAHVILACNNMARASEAVSRILEEWHKAKVETM 189  
QY 181 TLDLALLRSVQHFAFAKKNVPLHVCNNAATFALPWSLTKDGLTTFQVNHGHFYL 240  
DB 190 TLDLALLRSVQHFAFAKKNVPLHVCNNAATFALPWSLTKDGLTTFQVNHGHFYL 249  
QY 241 QLLQDVLCRSAPARVIVVSSSESHRFTDINDSLGKLDLFSRLSPTKNDYWMALAYNRSKLCN 300  
DB 250 QLLQDVLCRSAPARVIVVSSSESHRFTDINDSLGKLDLFSRLSPTKNDYWMALAYNRSKLCN 309  
QY 301 ILFSNELHRRILSPRGVTSNAVHPGNMYSNIHRSWVYTLFTLARPTKSMQCGAATTV 360  
DB 310 ILFSNELHRRILSPRGVTSNAVHPGNMYSNIHRSWVYTLFTLARPTKSMQCGAATTV 369  
QY 361 YCAAYPELEGLGMYFNCCRCMPSPQAQSEETARTLWALSERLIQERLGSQSG 414  
DB 370 YCAAYPELEGLGMYFNCCRCMPSPQAQSEETARTLWALSERLIQERLGSQSG 423

RESULT 8  
AAB84484  
ID AAB84484 standard; protein; 363 AA.  
XX AAB84484;  
AC AAB84484;  
XX  
XX 05-SEP-2001 (first entry)  
XX  
XX Amino acid sequence of FRA16D oxidoreductase (FOR) 1 gene cDNA.  
XX  
XX Cancer associated protein; FOR gene; FRA16D; fragile site; aphidicolin;  
KW chromosomal rearrangement; cancer; splice variant; DNA instability;  
XX FRA16D oxidoreductase; neoplasia.  
XX  
OS Homo sapiens.  
XX  
XX WO200144466-A1.  
PN  
XX  
XX 21-JUN-2001.  
PD

XX 15-DEC-2000; 2000WO-AU001539.  
XX PF  
XX 16-DEC-1999; 99AU-00004711.  
XX PR  
XX 19-APR-2000; 2000AU-00007025.  
XX PR  
XX (WOME-) WOMEN'S & CHILDREN'S HOSPITAL.  
XX  
XX Richards R, Ried K, Finniss M, Hobson L, Mangelsdorf M, Dayan S;  
PI Nancarrow J, Woollatt E, Baker E;  
XX  
XX WPI; 2001-398151/42.  
XX DR N-PSDB; AAH27866.  
XX  
XX Novel isolated 16q23.2 nucleic acid molecule, FRA16D oxidoreductase (FOR)  
XX gene associated with FRA16D site, useful for early diagnosis and  
XX PT assessment of risk of cancers associated with the FRA16D region.  
XX  
XX Claim 31; Fig 9; 150pp; English.  
XX  
XX

The present sequence represents a human FRA16D oxidoreductase (FOR) I  
CC transcript. The FOR gene encodes a cancer associated protein. The FRA16D  
CC site is a fragile site induced by aphidicolin, which is located within  
CC the FOR gene. The fragile site is the location of breakpoints of a  
CC variety of chromosomal rearrangements, and other mutations associated  
CC with cancers. The FOR protein is expressed as a number of splice  
CC variants. FOR gene polynucleotide fragments are capable of acting as  
CC specific primers or probes for detecting cancer associated variations of  
CC DNA sequence such as a point mutation or small DNA rearrangement  
CC associated with the tumour, a breakpoint of one or more chromosomal  
CC rearrangements associated with the tumour and a pause site within the  
CC FRA16 gene. FOR nucleic acid molecules are useful as markers to identify  
CC relationship between the fragile site (FRA16D) and the DNA instability in  
CC neoplasia which allows better diagnosis of cancers associated with the  
CC region  
XX

SQ Sequence 363 AA;

Query Match 84.8%; Score 1869; DB 4; Length 363;  
Best Local Similarity 99.7%; Pred. No. 1.1e-181;  
Matches 351; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAALRYAGLDGTTSEDELPDGGWEERTTKDGVVYANHTTEKTQWHPKTKGKRKRVAGDLP 60  
DB 1 MAALRYAGLDGTTSEDELPDGGWEERTTKDGVVYANHTTEKTQWHPKTKGKRKRVAGDLP 60  
QY 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDNPTKPTRQRYDGSITAMAILQGR 120  
DB 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDNPTKPTRQRYDGSITAMAILQGR 120  
QY 121 DFTGKVVVVTGANGSIGFETAKSFALHGAHVILACNNMARASEAVSRILEEWHKAKVETM 180  
DB 121 DFTGKVVVVTGANGSIGFETAKSFALHGAHVILACNNMARASEAVSRILEEWHKAKVETM 180  
QY 181 TLDLALLRSVQHFAFAKKNVPLHVCNNAATFALPWSLTKDGLTTFQVNHGHFYL 240  
DB 181 TLDLALLRSVQHFAFAKKNVPLHVCNNAATFALPWSLTKDGLTTFQVNHGHFYL 240  
QY 241 QLLQDVLCRSAPARVIVVSSSESHRFTDINDSLGKLDLFSRLSPTKNDYWMALAYNRSKLCN 300  
DB 241 QLLQDVLCRSAPARVIVVSSSESHRFTDINDSLGKLDLFSRLSPTKNDYWMALAYNRSKLCN 300  
QY 301 ILFSNELHRRILSPRGVTSNAVHPGNMYSNIHRSWVYTLFTLARPTKSM 352  
DB 301 ILFSNELHRRILSPRGVTSNAVHPGNMYSNIHRSWVYTLFTLARPTKSM 352  
RESULT 9  
AAB63100  
ID AAB63100 standard; protein; 409 AA.  
XX AAB63100;  
AC AAB63100;  
XX



Qy 121 DFTGKVVVVTGANGSIGFETAKSPFALHGAHVILACRNMARASEAVSRILEEW 172  
Db 121 DFTGKVVVVTGANGSIGFETAKSPFALHGAHVILACRNMARASEAVSRILEEW 172

RESULT 11  
ABBI0434  
ID ABB10434 standard; protein; 173 AA.  
AC ABB10434;  
XX  
XX  
DT 10-JAN-2002 (first entry)  
DE Human cDNA SEQ ID NO: 742.  
XX  
XX Human; gene therapy; neural disorder; immune system disorder;  
KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
KW pulmonary disorder; cardiovascular disorder; renal disorder;  
KW proliferative disorder; inflammation.  
XX  
OS Homo sapiens.  
XX WO200154474-A2.  
PN  
PN  
PD 02-AUG-2001.  
PF  
PF 17-JAN-2001; 2001WO-US001349.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-019123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-021680P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 23-AUG-2000; 2000US-0227182P.  
PR 30-AUG-2000; 2000US-0227009P.  
PR 01-SEP-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0228927P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.

PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254037P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX XX

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI, 2001-476161/51.

DR N-PSDB; ABA06656.

XX  
PT Isolated nucleic acid molecule encoding an inflammation-associated  
PT polypeptide is used in preventing, treating or ameliorating a medical  
PT condition.

XX Claim 11; SEQ ID NO 742; 859pp + Sequence Listing; English.

XX The present invention provides human cDNAs, proteins and related genomic  
XX DNAs. These can be used in the treatment of neural, immune system,  
XX muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,  
XX renal and proliferative disorders and inflammation. The present sequence  
XX is a protein of the invention

XX Sequence 173 AA;

Query Match 39.2%; Score 862.5; DB 4; Length 173;

Best Local Similarity 92.5%; Pred. No. 2.3e-79;

Matches 161; Conservative 2; Mismatches 10; Indels 1; Gaps 1;

QY 3 ALRYAGLDDTDSDELPPGMEERTTKDGVVYANHTTEKTQWEHPKTKKRVAGDLPYG 62

Db 1 ALRYAGLDDTDSDELPPGMEERTTKDGVVYANHTTEKTQWEHPKTKKRVAGDLPYG 60

QY 63 WEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDNPKPTTRQYDGSSTAMEILQGRDF 122

Db 61 WEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDNPKPTTRQYDGSSTAMEILQGRDX 120

QY 123 TGVVVVTVGANGSIGFTAKSFALHGAHVILACNMRASEAVSRILEEWHKAK 176

Db 121 TGVVVVTVGANGSIGFTAKSFALHGAHVILACNMRASEAVSRILEEWHKAK 173

RESULT 12

AAU18436

ID AAU18436 standard; protein; 173 AA.

XX AAU18436;

XX 21-NOV-2001 (first entry)

XX Human endocrine polypeptide SEQ ID No 391.

XX Endocrine protein; human; mouse; rabbit; goat; horse; food additive; cat;  
KW dog; chicken; sheep; immunosuppressive; antiarthritic; vasotropic;  
KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;  
KW cerebroprotective; neotropic; antibacterial; virucide; fungicide; cancer;  
KW ophthalmological; vulnary; gene therapy; autoimmune disease; neoplasm;  
KW hyperproliferative disorder; breast; liver; cardiovascular disorder;

KW cerebrovascular disorder; nervous system disorder; bacterial infection;  
KW fungal infection; viral infection; ocular disorder; endocrine disorder;  
KW gastrointestinal disorder; renal disorder; respiratory disorder;  
KW wound healing; skin aging; organ transplantation; food preservative;  
XX tissue regeneration; anti-infertility.

XX Homo sapiens.

XX WO200155364-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001308.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205151P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214886P.

XX 30-JUN-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216647P.

XX 07-JUL-2000; 2000US-0216800P.

XX 11-JUL-2000; 2000US-0217487P.

XX 11-JUL-2000; 2000US-0217496P.

XX 14-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

XX 26-JUL-2000; 2000US-0220964P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224519P.

XX 14-AUG-2000; 2000US-0225213P.

XX 14-AUG-2000; 2000US-0225214P.

XX 14-AUG-2000; 2000US-0225266P.

XX 14-AUG-2000; 2000US-0225267P.

XX 14-AUG-2000; 2000US-0225268P.

XX 14-AUG-2000; 2000US-0225270P.

XX 14-AUG-2000; 2000US-0225447P.

XX 14-AUG-2000; 2000US-0225757P.

XX 14-AUG-2000; 2000US-0225758P.

XX 14-AUG-2000; 2000US-0225759P.

XX 18-AUG-2000; 2000US-0226279P.

XX 22-AUG-2000; 2000US-0226681P.

XX 22-AUG-2000; 2000US-0226686P.

XX 22-AUG-2000; 2000US-0227182P.

XX 23-AUG-2000; 2000US-0227009P.

XX 30-AUG-2000; 2000US-0228924P.

XX 01-SEP-2000; 2000US-0229387P.

XX 01-SEP-2000; 2000US-0229343P.

XX 01-SEP-2000; 2000US-0229344P.

XX 01-SEP-2000; 2000US-0229345P.

XX 05-SEP-2000; 2000US-0229509P.

XX 05-SEP-2000; 2000US-0229513P.

XX 06-SEP-2000; 2000US-0230437P.

XX 06-SEP-2000; 2000US-0230438P.

XX 08-SEP-2000; 2000US-0231242P.

XX 08-SEP-2000; 2000US-0231243P.

XX 08-SEP-2000; 2000US-0231244P.

XX 08-SEP-2000; 2000US-0231413P.

XX 08-SEP-2000; 2000US-0231414P.

XX 08-SEP-2000; 2000US-0232080P.

XX 08-SEP-2000; 2000US-0232081P.

XX 12-SEP-2000; 2000US-0231968P.

XX 14-SEP-2000; 2000US-0232397P.

XX 14-SEP-2000; 2000US-0232398P.

XX 14-SEP-2000; 2000US-0232399P.

XX 14-SEP-2000; 2000US-0232400P.

XX 14-SEP-2000; 2000US-0232401P.

XX 14-SEP-2000; 2000US-0233063P.

XX 14-SEP-2000; 2000US-0233064P.



PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236803P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241783P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 01-DEC-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.

PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Barash SC, Ruben SM;  
XX MPI; 2001-451936/48.  
XX N-PSDB; AAS29665.  
XX Isolated polypeptide for treating, preventing and/ or prognosing  
PT disorders of the endocrine system such as reproductive disorders,  
PT endocrine cancers and also for testing and detection e.g. diagnosis.  
XX Claim 11; SEQ ID NO 391; 604pp; English.  
XX Sequences AAU18282-AAU18507 represent endocrine polypeptides of the  
CC invention. Endocrine polypeptides and their associated polynucleotides  
CC are useful in the diagnosis, treatment and prevention of various types of  
CC disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs,  
CC chickens or sheep. A pathological condition can be determined by  
CC determining the presence or absence of a mutation in an endocrine  
CC polynucleotide. The treatable disorders include autoimmune diseases such  
CC as rheumatoid arthritis, hyperproliferative disorders such as neoplasms  
CC of the breast or liver, cardiovascular disorders such as cardiac arrest,  
CC cerebrovascular disorders such as cerebral ischaemia, nervous system  
CC disorders such as Alzheimer's disease, infections caused by bacteria,  
CC viruses and fungi, ocular disorders such as corneal infection, endocrine  
CC disorders such as Crohn's disease, renal disorders such as  
CC glomerulonephritis and respiratory disorders such as asthma. The  
CC polypeptides can also be used to aid wound healing, to prevent skin aging  
CC tissues and in chemotaxis. The polypeptides can also be used as a food  
CC additive or preservative to increase or decrease storage capabilities.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences  
Query Match 39.2%; Score 862.5; DB 4; Length 173;  
Best Local Similarity 92.5%; Pred. No. 2.3e-79;  
Matches 161; Conservative 2; Mismatches 10; Indels 1; Gaps 1;  
QY 3 ALRYAGLDDTDSDELPGWEERTTKDQWVYANHTTEKTQWEHPTKGRKEVAGDLPYG 62  
Db 1 ALRYAGLDDTDSDELPGWEERTTKDQWVYANHTTEKTQWEHPTKGRKEVAGDLPYG 60  
QY 63 WEQETDENGQVFFVDHINKRTTYLDPLRAFTVDDNPTKPTTQRQYDGGTAMEILQGRDF 122  
Db 61 WEQETDENGQVFFVDHINKRTTYLDPLRAFTVDDNPTKPTTQRQYDGGTAMEILQGRDX 120  
QY 123 TGVVVVVTGANGSIGGFETAKSFALHGAHVILACRMARASEAVSILSEWHKAK 176  
Db 121 TGVVVVVTGANGSIGGFETAKSFALHGAHVILACRMARASEAVSILSEWHKAK 173  
RESULT 13  
ABP67021  
ID ABP67021 standard; protein; 173 AA.  
XX AC ABP67021;  
XX DT 09-DEC-2002 (first entry)  
XX DE Human polypeptide SEQ ID NO 742.  
XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;  
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;  
KW antiparkinsonian; antiscikling; antianaemic; antiarthritic; cancer;  
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antitubercular;





Db	76	GESAASEIRVDTKNSQVLRKLDLSDTKSIRAFAGFLAEKQLHILINNAGVMMCPYSK	135
QY	221	TXDGLTTTQVNHGHHFYLVQLQDVLCESAPARVIVSVSESHRFTDINDSLGKLDSESL	280
Db	136	TADGFETHLGVNHLGHFLTYLLLEKLVKSAPARVVNSVAHH-----IGKIFHDL	188
QY	281	SPTKNDYMWAMLAYNRSKLCNILLFSELHRRLSRPGVTSNAVHPGNMYSNIHRSWWVYTL	340
Db	189	QSEKR-YSRGFAYCHSKLANVLTRELAKRLQGTGVTYAVHFG-VVRSELVHSSLLCL	246
QY	341	LFTLARPFTKSMQOGAATTVYCAVPELEGLGGMVFNCCRCMPSPESAOSBETARTLWAL	400
Db	247	LWELFSPFVKTAREGAQTSIHCAEGLPLESGKYPSCDKRTWVSPARNNKTAERLWNV	306
QY	401	SEELI 405	
Db	307	CELL 311	

Search completed: May 24, 2004, 10:00:09  
Job time : 61 secs